TRAUNER VERLAG

UNIVERSITÄT

DOUCEK PETR ■ CHROUST GERHARD ■
OŠKRDAL VÁCLAV (EDITORS)

IDIMT-2020
Digitalized Economy, Society and Information Management

28th Interdisciplinary
Information Management Talks
Sept. 2–4, 2020
Kutná Hora, Czech Republic

JYU
JOHANNES KEPLER
UNIVERSITÄT LINZ

TRAUNER VERLAG

Kutná Hora, Italian Court

Schriftenreihe Informatik

49

Doucek Petr ■ Chroust Gerhard ■ Oškrdal Václav (Editors)

IDIMT-2020
Digitalized Economy, Society and Information Management

28th Interdisciplinary Information Management Talks
Sept. 2–4, 2020
Kutná Hora, Czech Republic

Impressum

Schriftenreihe Informatik

Doucek Petr ■ Chroust Gerhard ■ Oškrdal Václav
(Editors)
IDIMT-2020
Digitalized Economy, Society and
Information Management
28th Interdisciplinary Information Management Talks

Programme Committee
Aumayr Georg, AT
Buchalcevová Alena, CZ
Chroust Gerhard, AT
Dhungana Deepak, AT
Doležel Michal, CZ
Doucek Petr, CZ
Fischer Jakub, CZ
Gerhard Jahn, AT
Helfert Markus, IR
Maryška Miloš, CZ
Neubauer Georg, AT
Novotný Ota, CZ
Oškrdal Václav, CZ
Pavlíček Antonín, CZ
Pitner Tomáš, CZ
Pucihar Andreja, SI
Purcarea Anca Alexandra, RO
Rainer Karin, AT
Sonntag Michael, AT
Schoitsch Erwin, AT
Tkáč Michal, SK
Wagner Jaroslav, CZ
Yablotchnikov Sergej, RU
Zimmermann Hans-Dieter, CH

Table of Contents

TOWARDS 2030 – DIGITAL TRANSFORMATION IN CRISIS MANAGEMENT

SOCIAL MEDIA, PRIVACY AND ETHICS

SMART SUPPLY CHAIN

CYBER SECURITY AND DIGITAL

SUSTAINABILITY AND PERFORMANCE MANAGEMENT

DIGITAL SECURITY, SAFETY, TRUST OF SMART SYSTEMS – FOR A RESILIENT

PEOPLE AND PROCESSES IN ENTERPRISE SOFTWARE DEVELOPMENT

HUMAN AND CULTURAL IMPACTS ON PROJECT SUCCESS

ANNEX

A Hearty Welcome to IDIMT 2020!

This year we are meeting again in the historic town of Kutná Hora, which demonstrates the resilience and preservation of old values and perspectives. The municipal administration of Kutná Hora also stresses continuity from the past by restoring places and streets in the old style (e.g. re-introducing cobblestones). Kutná Hora is the 7th location hosting the IDIMT Conferences (see Fig. 2 below).

Fig 1: Kutná Hora, Italian Court

In this year, 2020, however, we live under dramatic differences to the previous years. The Corona virus has not as yet been defeated. It has demonstrated the vulnerability and fragility of our high-tech society with respect to human health and the resulting economic and social situations. Initially we feared that we might not be able to hold a face-to-face conference at all and considered virtual alternatives – but as a consequence virtual inclusion of participants has been planned for.

The world-wide challenges posed by the Corona virus have demonstrated to us the importance of Interdisciplinarity of approaches and the need for Management of Information – two topics which have accompanied the IDIMT conferences since their beginning in 1993! [1]

The IDIMT conferences have shown a remarkable growth from 13 participants in 1993 to around 100 in 2020. From the start in 1993 the over-all focus of the IDIMT conferences - as expressed in the name – has always been Interdisciplinarity and Information Management. The specific aspects of these topics as discussed during the 27 previous conferences have changed according to the needs of the participants and the evolution of the world around us. Crisis management is a hot topic this year.

This year we have received 64 submitted papers plus 11 keynote papers and 1 invited papers with a total of more than 110 co-authors. Theses submissions have been reviewed in a two-step submission process with a double blind review process in order to judge and decide their acceptability. The authors have received extensive comments and have been encouraged to improve their papers by use of these. This often enabled us to promote their status from 'rejected' to 'accepted'.

The authors have come from 12 different countries: Austria, Czech Republic, Estonia, Germany, India, Kazakhstan, Lithuania, Netherlands, Poland, Russia, Slovakia and Ukraine. The programme will run in two parallel streams. In the conference we expect approximately 90 participants.

For 2020 we have chosen the following topics (arranged in approximately decreasing number of submissions)

- Innovation, New Business Models and Strategies
- Social Media, Privacy and Ethics
- Digital Economy and Industry 4.0
- Sustainability and Performance Management
- Smart Supply Chain
- Towards 2030 – Digital Transformation in Crisis Management
- Human and Cultural Impacts on Project Success

---

[1] Chroust, G. and Doucek, P., editors (1993). Information Management Workshop 93, Kubova Hut, Czech Republic. Univ. of Economics Prague & J. Kepler University Linz, Austria 1993, ISBN 3-902457-06-6

- Digital Security, Safety, Trust of Smart Systems – for a resilient Society

- People and Processes in Enterprise Software Development

- Cyber Security and Digital Identities

Most of the topics have been carried over from the previous conferences, sometime with some change of focus. We are also pleased to welcome two new topics (‚Human and Cultural Impacts…‘ and ‚People and Processes… .').

The distribution of the papers sheds some light on the current interests of our participants: the first two topics drew almost 40% of the submissions, the first 4 topics together amount to 60% of the submissions. The last 4 topics together drew only 20 % submissions.

Each session is organized by a Session Chairperson and traditionally starts with a keynote, followed by papers providing additional points of view. At the end of each session there is a 20 minute, often heated, discussion. This extensive interdisciplinary exchange of thoughts is one of the unique features of the IDIMT-Conferences.

Since 2000 Christian Loesch has always offered a special contribution: an overview of technical, economic and/or business developments in the global world.

To preserve the impressive history of the IDIMT conferences we have separately published a hard cover book and its e-book version[2]. Additionally we have collected Christian Loesch's 21 presentations (2000--2020!) and republished them in a separate book[3].
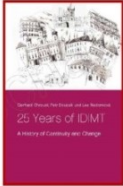
Another traditional highlight of IDIMT is the afternoon/evening excursion, excellently organized by Petr Doucek and his team. It has always offered a visit to a culturally impressive sights followed by a hearty and very plentiful dinner in a typical Czech inn. This year Kutná Hora itself provides the cultural background and gives the participants the chance to appreciate the historical importance of this city.

Read about the history of IDIMT!
IDIMT Memorial Publications
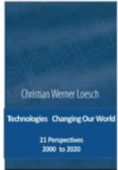
CHROUST G., NEDOMOVÁ, L., DOUCEK, P. (eds.):
25 Years of IDIMT:
A History of Continuity and Change
Books on Demand, Norderstedt, Germany, 2017 (hard copy and e-book), 60 pages
ISBN 978-3-74480-9573
Price: hardcopy 14 €, e-book 4,50 €

LOESCH, C.W., (ed.):
Technologies Changing Our World
21 Perspectives 2000 to 2020
Books on Demand, Norderstedt, Germany, 2020 (hard copy and e-book), 288 pages.
ISBN 978-3-75196-7693
Price: hardcopy 20 €, e-book ca. 7 €

Order information: books can be ordered at : BoD Buchshop
https://www.bod.de/buchshop/ and in international bookshops like Amazon.

The preparation and realization of IDIMT 2020 would not have been possible without the support of many organizations and persons. Therefore we would like to express our thanks to:

- the University of Economics, Prague for the project IGA 409039,

- the Faculty of Informatics and Statistics of the University of Economics, Prague, and

- the Johannes Kepler University Linz, Institute of Telecooperation.

Our further thanks go to:

- Petr Doucek for chairing the Organizing Committee, for arranging the conference location, the hotels and the greatly appreciated evening event,

- Gerhard Chroust for chairing the Programme Committee,

- Antonín Pavlíček and Lea Nedomová, for organizing the program, the reviews, keeping contact with all involved speakers, and reminding forgetful authors and session chairs,

- Václav Oškrdal for arranging and assembling the selected papers for the proceedings,

- Lea Nedomová, for her support in performing the necessary administrative tasks,

- all Keynote Speakers, speakers and contributors of papers,

- all members of the Programme committee and the Session Chairpersons for soliciting contributors and creating an interesting and compact program,

- all reviewers providing critical remarks for improving the papers,

- the Trauner Verlag for acting as the publisher of our conference, and

- all other unnamed persons contributing to the success of this conference.

Looking forward to a successful and interesting conference!

Gerhard Chroust   July 2020   Petr Doucek

The seven locations of IDIMT Conferences 1993 – 2020

Kubova Huť	1993, 1994
Zadov	1995 – 2002
České Budějovice	2003 – 2007
Jindřichův Hradec	2008 – 2012
Praha	2013
Poděbrady	2014 – 2017
Kutná Hora	2018 – 2020

Fig. 2: The 7 locations of IDIMT Conferences

A very short history of Kutná Hora

Kutná Hora, as its name suggests, was originally a thriving silver miners' town (*kutání* in medieval Czech means 'mining)'. In 1142 the first Cistercian monastery in the Czech lands was established in the nearby village of Sedlec and at the end of the 13th century the original mining settlement of Cuthna antiqua – Old Kutna - was founded. The city was located above a high deposits of silver ore and was a successful mining town. It was the strongly fortified 'treasury' of the Bohemian kings. As a consequence it became Bohemia's political and economic center between the 14th and 15th century and once was one of the richest cities of Bohemia. The Barbara's Cathedral was founded ('privately'!) by one rich mine owner (!) in 1388. Today the silver mines are depleted, but still can be visited. A visit is planned for the IDIMT conference!

Sponsors of IDIMT 2020

General partner / Generální partner VŠE v Praze

INVITED CONTRIBUTION

TECHNOLOGIES CHANGING OUR WORLD

Christian W. Loesch

IBM ret.

CWL001@gmx.net

Keywords

ICT industry and economy, future of microelectronics, more Moore and beyond Moore, emerging technologies, quantum, neurocomputing, sensors.

Abstract

Based on an analysis of the economic status of the ICT industry, we will peruse the present status and future developments of microelectronics from "more Moore" to" beyond Moore". On the threshold of new computing paradigms we will look at emerging technologies, progressively important areas as communication and sensor technology as well as the arising challenges and problems.

1. Economic Scenario

ICT industry has changed dramatically in the last few years, with 2019 being a turnaround year as shown by the economic developments of some key players of the industry below.
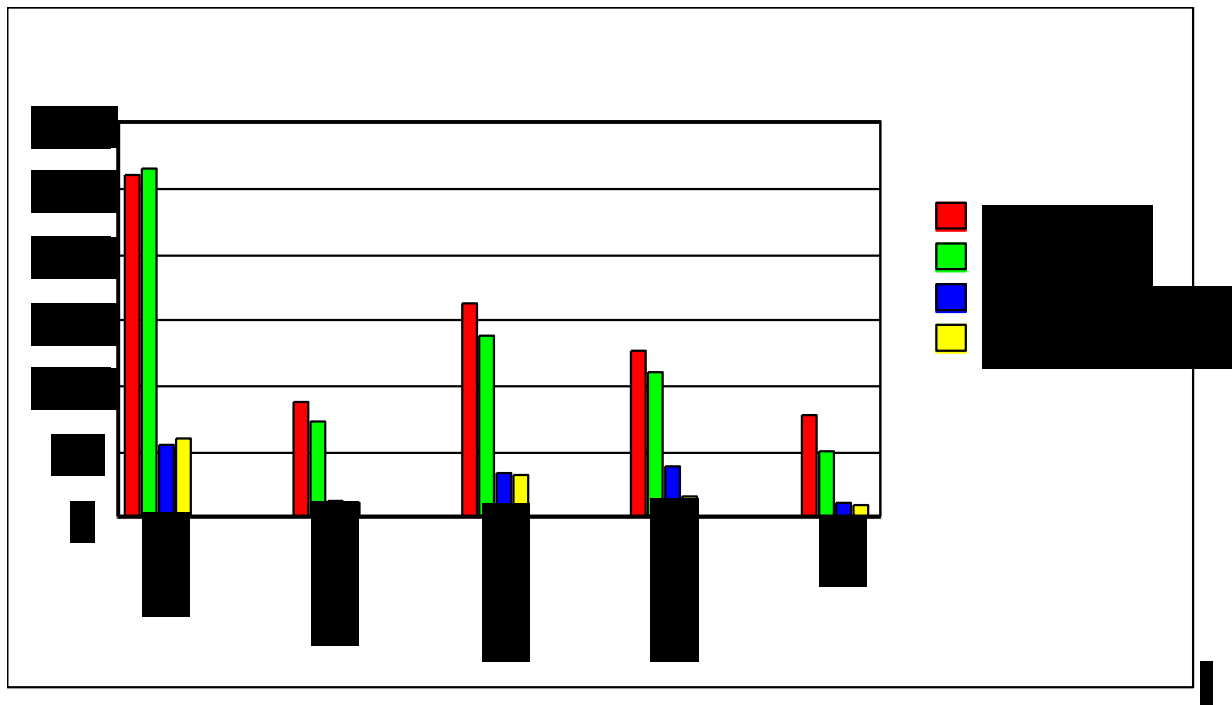
How did key players fare 2019

How and where are they achieving their impressive results:

Apple	- 4%
Amazon	+11%
Alphabet	+18%
Microsoft	+14%
IBM	+48%

- Apple: Diversifying
- Alphabet: Google, YouTube, etc. Adv.> 85% of rev.
- Microsoft: Most diversified, tax, five divisions each 20 %
- IBM: Not comparable due to new accounting standards

The worldwide market for chips has reached in 2017 the impressive volume of 412 b$ representing a rise of 21,6%. The IC market forecast (by IC Insights) for 2020 expected strong growth again of 8,0% and units shipments up 7,0%. In parallel a concentration process has reduced the number of leading edge chip manufacturing companies from 28 in 2001 to 5 in 2018.

Let's hope that these successes have been used by the industry to build the resilience needed to overcome the events of 2020.

1Q20 Top 10 Semiconductor Sales Leaders ($M, Including Foundries)

1Q20 Rank	1Q19 Rank	Company	Headquarters	1Q19 Total IC	1Q19 Total O-S-D	1Q19 Total Semi	1Q20 Total IC	1Q20 Total O-S-D	1Q20 Total Semi	1Q20/1Q19 % Change
1	1	Intel	U.S.	15,799	0	15,799	19,508	0	19,508	23%
2	2	Samsung	South Korea	11,992	875	12,867	13,939	858	14,797	15%
3	3	TSMC (1)	Taiwan	7,096	0	7,096	10,319	0	10,319	45%
4	4	SK Hynix	South Korea	5,903	120	6,023	5,829	210	6,039	0%
5	5	Micron	U.S.	5,465	0	5,465	4,795	0	4,795	-12%
6	6	Broadcom Inc. (2)	U.S.	3,764	419	4,183	3,700	410	4,110	-2%
7	7	Qualcomm (2)	U.S.	3,753	0	3,753	4,050	0	4,050	8%
8	8	TI	U.S.	3,199	208	3,407	2,974	190	3,164	-7%
9	11	Nvidia (2)	U.S.	2,215	0	2,215	3,035	0	3,035	37%
10	15	HiSilicon (2)	China	1,735	0	1,735	2,670	0	2,670	54%
–	–	Top-10 Total		60,921	1,622	62,543	70,819	1,668	72,487	16%

(1) Foundry   (2) Fabless

Source: Company reports, IC Insights' Strategic Reviews database

But the events of 2020 are changing the previous assumptions dramatically as shown below.

How Covid-19 Outbreak Altered 2020 Outlook

IC Total $384.8B 8% Growth — $14.2B → **IC Total $370.6B 3% Growth** — $24.8B → **IC Total $345.8B 4% Decline**

Microcomponents
Memory
Logic
Analog

MR20 (Jan. '20) | **March Update** | **April Update**

Source: IC Insights
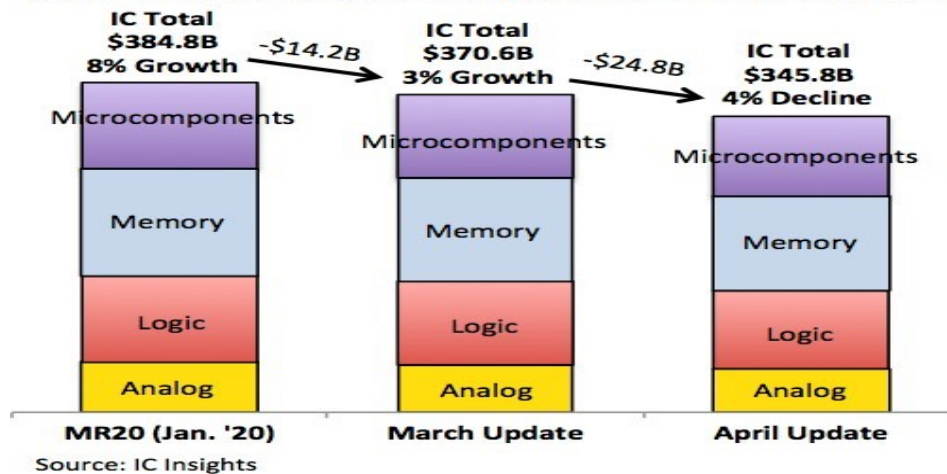
2. Technology

The twilight of Moore's law does not mean the end of progress. Innovation will continue, but it will be more sophisticated and complicated. Remember what happened to airplanes? A Boeing 787 doesn't go faster than a 707 did in the 1950s, but they are very different airplanes, with innovations ranging from fully electronic controls to a carbon-fiber fuselage. That may happen with computers. New EUV scanners will expand Moore's law for the anticipatable future, but nobody should overlook the equally impressive tacit advances in performance, e.g. the current Intel Core i3 processor is 32% faster than the first top of the line Intel Core i7 at half the power consumption only.

The chip making process is getting exceedingly complex, often involving hundreds of stages, meaning that taking the next step down in scale requires a closely intertwined network of materials suppliers and apparatus developers and manufacturers etc. to deliver the right new developments at the right time. If you need 40 kinds of equipment and only 39 are ready, then everything stops.

Leading companies, are trying to shrink components until they limits of the wall of quantum effects. The more we shrink, the more it costs. Every time the scale is halved, manufacturers need a whole new generation of ever more precise photolithography machines. Building a new fab line today requires an investment typically measured in billions of dollars, an investment only few companies can risk and afford this like INTEL, GLOBALFOUNDRIES, Samsung or TSMC. All of these companies rely on high volume manufacturing to finance the capital and the enormous R&D requirements to maintain their competitiveness,

Worlwide Wafer Capacity Leaders
(Monthly Installed Capacity in Dec 2019, 200mm-equivalents)

2019 Rank	2018 Rank	Company	Headquarters Region	Dec-2018 Capacity (K w/m)	Dec-2019 Capacity (K w/m)	Yr/Yr Change	Share of Worldwide Total	Inclusion or Exclusion of Capacity Shares from JV Fabs
1	1	Samsung	South Korea	2,934	2,935	0%	15.0%	
2	2	TSMC	Taiwan	2,439	2,505	3%	12.8%	shares of SSMC & VIS
3	3	Micron	North America	1,685	1,841	9%	9.4%	share of IM Flash in '18
4	4	SK Hynix	South Korea	1,630	1,743	7%	8.9%	
5	5	Kioxia/WD	Japan	1,361	1,406	3%	7.2%	

Source: Companies, IC Insights' *Global Wafer Capacity 2020-2024 Report*

The old market was characterized by producing a few different products, selling large quantities of them. The new market is producing a huge variety of products, but selling a few hundred thousand apiece, so costs of design and production has to be low. The fragmentation of the market triggered by mobile devices is making it additionally harder to recoup the investments. As soon as the cost per transistor at the next node exceeds the existing cost, the scaling stops. We may run out of money before we run out of physics.

Computing is increasingly defined by high-end smartphones, tablets, and other wearables, as well as by the exploding number of smart devices everywhere from bridges to the human body. These mobile devices have requirements different from their more sedentary cousins. The chips in a typical smartphone must send and receive signals for voice calls, Wi-Fi, Bluetooth and GPS, while also sensing touch, proximity, acceleration, magnetic fields, even fingerprints, demanding the device to host special purpose circuits. In this form the user value doubles every two years, Moore's law will continue as long as the industry can keep successfully marketing devices with new functionality.

Advanced Digital Computing (More Moore)

As shown below leading companies expect Moore to continue for years. Digital CMOS is currently at the 14 nm node with potential to scale to 3 nm by 2022. The challenges are materials and process variation to achieve these with new technology at acceptable tool and fabrication costs.

2.1. Emerging technologies and paradigms.

We are on the threshold of revolutionary new computing paradigms. We can look forward to a decade of multiple technologies going to revolutionize the world of computing over the next 5-10 years.

Over the last decades, intensive efforts have been made on enhancing the capabilities and performance potential of III-V wide bandgap material systems such as Indium Phosphide, Gallium Arsenide, Silicon Germanium, Silicon Carbide, Gallium Nitride, and Aluminum Nitride.

Parallel to this evolves the architectural approach: stick with silicon, but configure it in new ways to using 3D to pack more computational power into the same space. 3D sequential integration is an alternative to conventional device scaling. Compared to TSV-based 3D ICs, 3D sequential process flow offers the possibility to stack devices with a lithographic alignment precision (few nm) enabling a density >100 million/mm^2 between transistors tiers (for 14nm), to merge several technologies and materials with 3D sequential integration of various devices.

However, this rather works with memory chips, which do not have the thermal problem as they use circuits consuming power only when a memory cell is accessed.

We will also address some farther out are options and paradigmata like quantum computing, or neuromorphic computing. But most of these alternative paradigms has made it very far out of the laboratory.

Compound Semiconductor

Over the last several decades, industry, academia and government have collaborated to deliver the enhanced capabilities and performance potential of III-V wide bandgap material systems such as Indium Phosphide, Gallium Arsenide, Silicon Germanium, Silicon Carbide, Gallium Nitride, and Aluminum Nitride as well as recent work on ultra-wide bandgap compound semiconductors, subsystem and system levels.

Despite the potential for enhanced performance of III-V compound semiconductors, it has not been generally adopted for integration into consumer products. This is due to material complexity, high cost and a lack of requirements for the high power and advanced capability offered.

However, certain sectors in the commercial market have transitioned to compound semiconductor technology replacing silicon technology, specifically in wireless mobile communication infrastructure (base stations), CATV, IoT, automotive and energy sectors. As availability of compound semiconductor material continues to grow, specifically GaN/SiC, costs will decrease and integration into consumers' systems will gain popularity.

Emerging technologies that may radically change the IT scenario are paradigms that diverge from simple transistor based logic and operations. Advanced Research is apparent for spintronic majority gate technologies including spin-based logic, graphene-based Tunneling Field Effect Transistor (TFET) technology and novel material FETs technology.

The evolution of transistor architecture and channel materials (MOSFET)
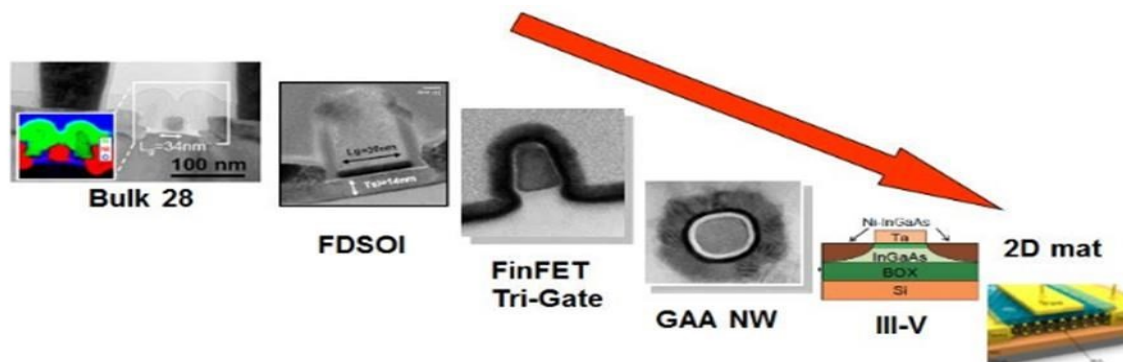
G. Ghibaubo, CNRS Grenoble

What makes these Nanotechnologies so appealing?

Remember: Carbon nanotubes (CNTs) are hollow cylinders composed of one or more concentric layers of carbon atoms in a honeycomb lattice arrangement, with a typical diameter of 1-2nm. Depending on the arrangement of the carbon atoms, the CNTs can be either metallic or semiconducting, and are considered both for interconnect or as field effect transistors (FETs).

The expected benefits of FETs over Silicon based devices are:

- High mobility is very high in carbon nanotubes, significantly higher than in any other material, enabling higher speed, or reduction of the operating voltage and lower active power (heat).

- The tube diameter is controlled by chemistry not by printing, allowing to reduce the body dimension beyond what is achievable lithography. This allows the fabrication of aligned arrays with high packing density.

- The intrinsic capacitance is a quantum capacitance related to the density of states and independent of electrostatics. The device capacitance could hence be much lower than the FinFETs gate to channel capacitances, reducing the switching energy.

Ferroelectric semiconductors and two-dimensional devices

Engineers at Purdue University and Georgia Tech constructed devices from a new two-dimensional material that combines memory-retaining properties and semiconductor properties using a newly developed ferroelectric semiconductor, alpha indium selenide. Noticeable applications would be: a type of transistor that stores memory as the amount of amplification it produces; and a two-terminal device that could act as a component in future computers using neuromorphic low-power AI chips as memristors as the neural synapses in their networks. Under the influence of an electric field, the molecule undergoes a structural change that holds the polarization. Even better, the material is ferroelectric even as a single-molecule layer only about a nanometer thick.

Digital reality, cognitive technologies, and blockchain are growing fast in importance. Virtual reality and augmented reality are redefining the fundamental ways humans interact with their surroundings, with data, and with each other. Cognitive technologies such as machine learning, robotic process automation, natural language processing, neural nets, and AI moved from highly special capabilities to tenets of strategy. These trends are poised to become as familiar and impactful as cloud, analytics, and digital experience are today.

Future memory technologies

MRAM has advantages over other memory technologies. Reading and writing data can be done at speeds similar to volatile technologies but consumes less power and, is nonvolatile, does not need a steady power supply to retain data.

MRAM stores information as the spins of electrons—a property related to an electron's intrinsic angular momentum. Most electrons in a ferromagnet point in the same direction. A currents magnetic field can cause most of those electrons to change their spins. The magnet records a "1" or a "0" depending on which direction they point.

But ferromagnets can be influenced by external magnetic fields, and the spins of adjacent ferromagnets can influence one another requesting enough space between them, limiting MRAM's ability to scale to higher densities for lower costs.
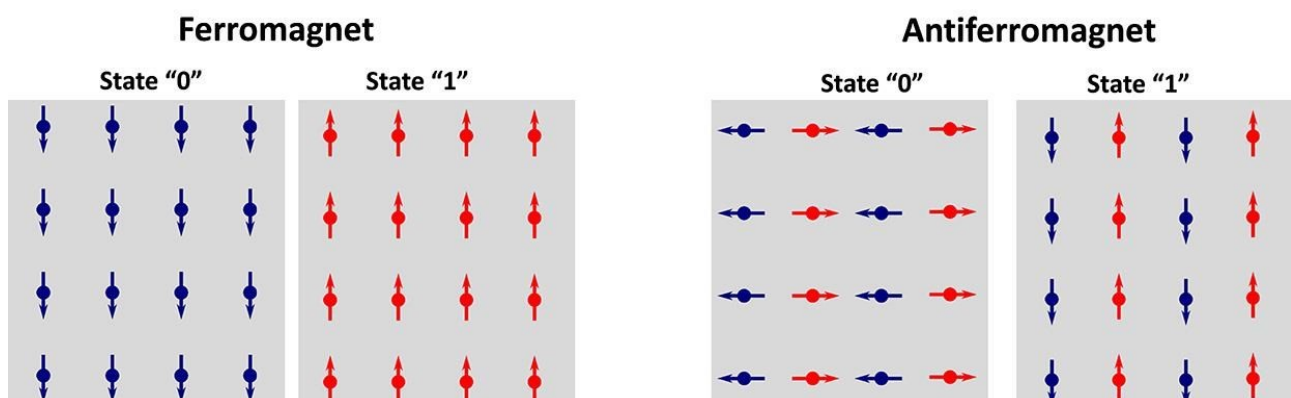
Pete Wadley

Ferromagnets [left] and antiferromagnets [right] can both store information in the spins of their electrons. But the orientations of those spins and their magnetic moments cancel out in antiferromagnets, making them impervious to external magnetic fields.

Antiferromagnets (metals such as Mn, Pl, Sn) do not have that problem. Electrons on neighboring atoms point opposite to each other and due to the dynamics of the spin in antiferromagnets are much faster, bits can be switched in picoseconds with terahertz frequencies. Theoretically, antiferromagnets could increase the writing speed of MRAM by three orders of magnitude. .

Analog Computing and Neuro-inspired computing

- Analog, neuromorphic and quantum computing paradigms each involve alternative gate sets and architectures facilitating new computing paradigms. However new computing paradigms will also create additional security challenges beyond the ones already present with advanced CMOS. Current interests focus on machine learning and AI enabling applications, and the search for the hardware implementations.

- Analog computing is receiving increasing attention with advanced SiGe RF technology, hybrid digital/analog platforms, NEMs, photonics and superconducting electronics. This paradigm is particularly well suited for sensor applications and has significant power advantages for certain other applications as well.

- Neuromorphic and Neuro-inspired computing is experiencing rapid growth with major companies having intensive R&D efforts in this area (Google, Amazon, IBM, Microsoft etc.).

- The present digital technology falls short, partly because device scaling gains are no longer easy to come by, and the intractable energy costs of computation. Deep learning, using labeled data, can be mapped onto artificial neural networks, arrays where the inputs and outputs are connected by programmable weights, which can perform pattern recognition functions. The learning process consists of finding the optimum weights, however this learning process is very slow for large problems. Exploiting the fact that weights do not need to be determined with high precision, research has recognized that analog computation approaches, using physical arrays of memristor (programmable resistor) type devices could offer significant speedup and power advantages compared to pure digital, or pure software approaches

- Machine vision - Machine vision technology has made great progress in recent years, and is now becoming an integral part of various intelligent systems, including autonomous vehicles and robotics. Usually, visual information is captured by a frame-based camera, converted into a digital format and processed afterwards using a machine-learning algorithm such as an artificial neural network (ANN). A large amount of (mostly redundant) data passes through the entire signal chain, however, results in low frame rates and high power consumption. Various visual data preprocessing techniques have thus been developed to increase the efficiency of the subsequent signal processing in an ANN demonstrating that an image sensor can itself constitute an ANN that can simultaneously sense and process optical images without latency. L. Mennel and his team (TU Vienna) demonstrated trained sensors to classify and encode images optically projected onto the chip with a dramatically increased throughput.

Impressed by these technological advances we have to keep in mind that most have evolved yet past the phase of a lab prototypes. The challenge may be 3D integration at affordable cost making organic materials an attractive candidate.

3. From Electronics to Photonics

Silicon photonics for optical quantum technologies is both technological as well as economically highly attractive. A fast expanding market both long-term with a CAGR 78–20 Fc of 8,6% and an accelerating growth rate in the last ten years up to 100%.This results in a continuous emphasis on future investment in R&D.(Statistics 2017).

Modern silicon photonics opens new possibilities for high-performance quantum information processing, such as quantum simulation and high-speed quantum cryptography.

- Solid state quantum memories based on electronic and nuclear spins are now becoming competitive for quantum repeater networks and distributed quantum computing

- Opto-electronic devices and 2D materials

2D materials, such as graphene, provide new capabilities in communications, sensing, imaging, nonlinear optics, and quantum information devices. There are theoretically about 16000 materials are eligible as candidates for single or combined 2D i.e. multilayer materials.

- Silicon Lasers

Silicon is the dominating and most thoroughly investigated material of microelectronics, seems to have another encouraging surprise ready. Emitting light from silicon has been the 'Holy Grail' in the microelectronics industry for decades.

Current technology, based on electronic chips, is reaching its ceiling. A limiting factor being heat, resulting from the resistance that electrons experience when traveling through the copper lines connecting the many transistors on a chip. To continue transferring more and more data, we need a new technique that does not produce heat as photonics.

In contrast to electrons, photons do not experience resistance. As they have no mass or charge, they will scatter less within the material they travel through, and therefore no heat is produced. The energy consumption will therefore be reduced. Moreover, by replacing electrical communication within a chip by optical communication, the speed of on-chip and chip-tochip communication can be increased by a factor 1000. Data centers would benefit especially, with faster data transfer and less energy usage for their cooling system. But these photonic chips will also bring new applications within reach. Think of laser-based radar for self-driving cars and chemical sensors for medical diagnosis or for measuring air and food quality

Silicon's mature and large-scale manufacturing base could lead to implement a much needed reduction in the cost of photonic devices. Such a cost reduction can bring the power of optical networks to the desktop computer and to home systems. It could enable a new generation of electro-opto-mechanical chips that perform the job of today's complex systems at a fraction of the cost, size, and power dissipation.

Let us make a short review of the basic principles to explain the problem.
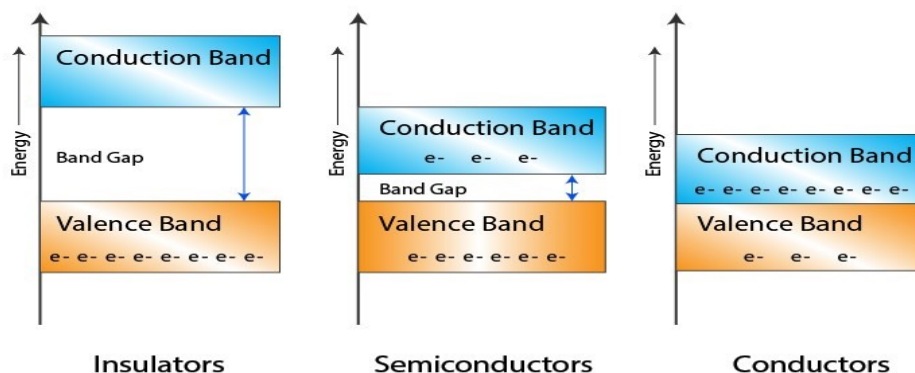
 But reality is more sophisticated because unfortunately for our purposes, direct band gap light emission is necessary whereas Si has the property of only indirect bandgap emission.
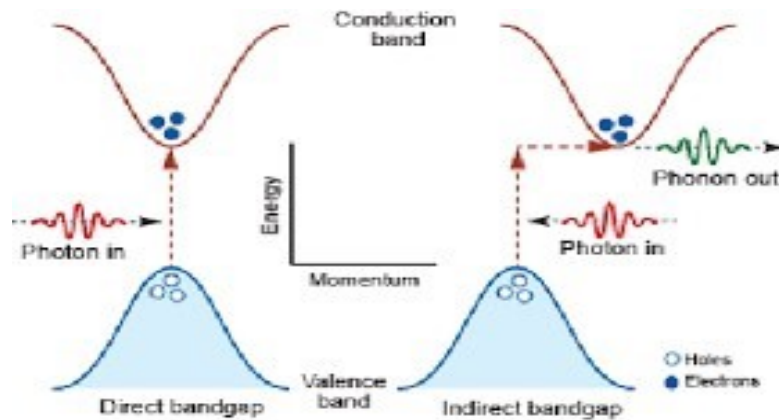
Unfortunately indirect bandgap semiconductors are usually very inefficient emitters. This problem been approached and resolved by an unusual approach. Researchers from TU of Eindhoven developed an alloy with silicon that has the desired properties to emit light and are now starting to create a silicon laser to be integrated into current chips.

Since QC and AI and related subjects have been covered in preceding IDIMT sessions only some additional comments:

Quantum-enhanced sensing

Quantum sensors enable unpaired precision measurements of time, fields, and forces for applications in the physical and life sciences.

QC (Quantum Computing)

QC continues to be perused with remarkable R&D (and PR) efforts to take advantage of the large parallelisms possible for complex optimization and factoring problems. It will not replace conventional computing but potentially offer superior performance for specific niche applications, rather than for the everyday digital computing tasks.

AI

AI is showing an impressive development. Factors responsible of its triumphal march are: more data, cheaper storage capacities and higher computing power (e.g. graphics card farms). They enable the use of AI processes in increasingly complex configurations. Experts differentiate between "strong AI", aiming to imitate human intelligence and "weak AI", which is used to make intelligent decisions for specific areas, such as the automation of processes, but strong AI is yet beyond the current technical possibilities. Unresolved fundamental problems ensure that it remains a theoretical game of thought for the foreseeable future, even if some of the reporting suggests otherwise. Weak AI, on the other hand, is an approach that plays a role in many applications today.

Further out in the long range future are "wet" technologies as the

Molecular computer

French scientists have built the first molecular computer using polymers to store data. They encoded and read the word the word "Sequence" in ASCII code using a synthetic polymer sequence, thus proving that it is possible to store information in polymer molecules. Given the size of each monomer unit of the molecule, this method would make the storage required for of each bit of information, a hundred times smaller than that of current hard drives.
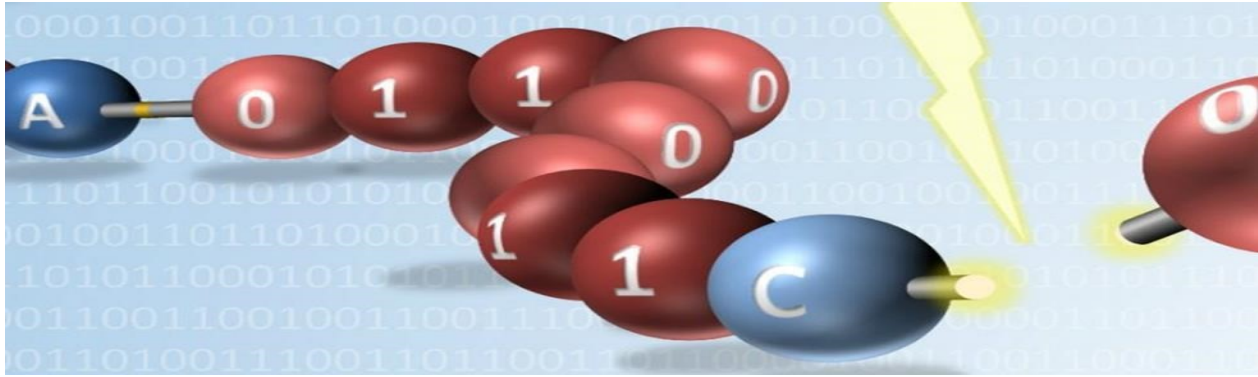
3.1. Lateral challenges, problems and risks emerging

New computing paradigms will create new security challenges. Analog computing, neuromorphic computing and quantum computing paradigms each involve alternative gate sets and architectures. The advancement of such emerging technologies will likely outpace industry's ability to understand the related security threats as well as the readiness of adequate legislation.

Ecology is another important aspect i.e. finding alternatives for rare or toxic materials, and processes.

4. Communication (Connectivity and Advanced Logic)

Networking has lived in the shadow of the high profile technologies but this is changing even more drastically than forecasted Communication is overtaking the computer IC market segment already and is expected to race ahead of all other end-uses (2020 McClean Report).

The connectivity functions will be everywhere in the connected world, from the physical world, (things and persons, autonomous objects, (factory 4.0, autonomous vehicles...), the Cyber Physical Systems), cloud, (E-Health, Intelligent Transport Systems, E-Security, E-Functions and computing).

The coming global skin of thousands of additional satellites will impact the scenario dramatically.

Expectations for the next 5 years (Source: Statistics 2017 but these figures may be dramatically impacted by Covid 19.
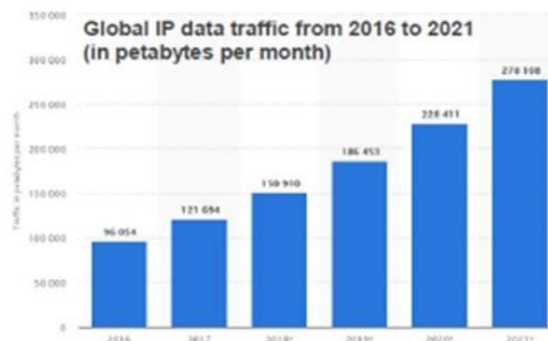
- Traffic   x 10000
- Energy/ bit   1/ 1000
- Connection density   x 10
- Data Rate/connection average   100Mb/s

- Data Rate/connection peak   10Gbs

- Latency   < 1ms

- Connected devices (w/o IoT)   x 100

- IoT connections   > 100 room

Technologies for Wireless applications (Indoor):

CMOS technology processes are the main circuit integration technology used and will continue the prevailing technology for the years to come. Challenges in this field will be the 3D integration in one low cost package. The best candidate for the economic point of view are organic materials, if they can demonstrate their technical capacity.

But the euphoric views and PR of some are met by the market participants with some skepticism. As the overhyped trends of 2020 have been quoted:

1. Augmented reality (39%)

2. 5G wireless technology (35%)

3. 3: Biometric authentication (32%)

4. AI in the data center (31%)

5. Blockchain (31%)

6. Anything "as a service" (30%)

A astonishing approach for optical communication technology in the long run may come with Vortex Lasers

Light has several degrees of freedom (wavelength, polarization, pulse length, and so on) that can be used to encode information. A light beam or pulse can also be structured to have the property of orbital angular momentum, becoming a vortex. Because the winding number of the vortex can be arbitrary, this technology opens the possibility to expand the channel capacity considerably.
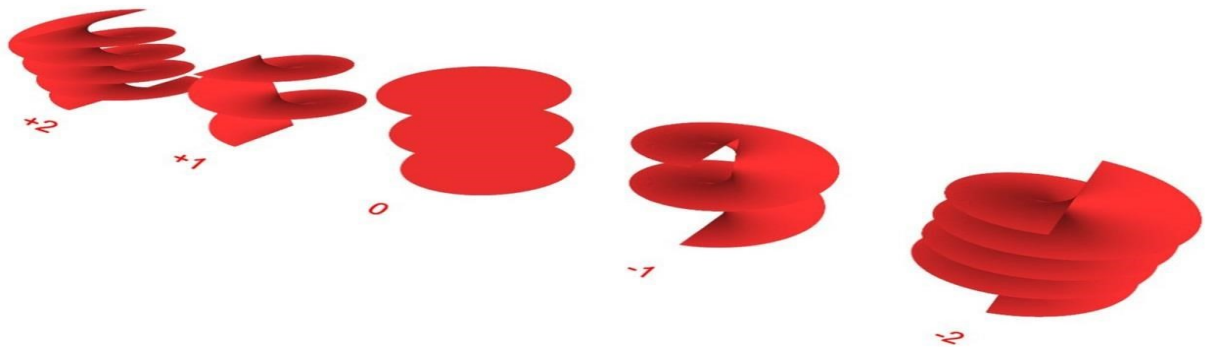
5. Sensors

Global endeavors aim at more sustainable, ICT-enabled strategies for healthcare, energy and environment. Overall, connected objects, IoT, big data, software and algorithms, zero-power or selfpowered sensors, sensor fusion, wireless sensor networks and system-in-package are all important for a future scenario. Improvements in healthcare sensor technology could drive an economic benefit of healthcare costs. Most of the sensors types mentioned below are similar and relevant for other industrial segments such as consumer electronics (MEMS accelerometers, magnetic, chemical and

gyroscopes), industrial (image sensors), and environment (air quality gas sensors) and defense (LiDAR sensors).

This may enable a plethora of applications in the fields of energy and environment as:

- Automotive:
  - The road transport sector should be 50% more efficient by 2030
  - CO_2 emissions will reduce significantly (80% cars, 40% trucks)
  - Transport schedules (mobility) will be more reliable and traffic safety will improve. Industry expects autonomous cars to improve safety of passengers and pedestrians, reducing fuel consumption by 10% and cost of insurance by 30%.
- Sensors for internal system performance: Motion, Pressure and Position, Advanced Driver Assistance System (ADAS)
- Image (recognition), LiDAR and Infrared sensors
- Environmental monitoring
  - Gas and Particulate sensors
- Medical:
  - Physiological signal monitoring
  - Implantable sensors
  - Molecular diagnostics
  - Telemedicine (analyst and diagnostic systems).
- Quantum enhanced sensing -Quantum Radar
  - An emerging remote-sensing technology based on quantum correlations (quantum entanglement) and output quantum detection, will allow the radar system to pick out its own signal even when swamped by background noise. This would allow to detect stealth aircraft, filter out jamming attempts, and operate in areas of high background noise.

6. Summary

The preceding review based on pre-Covid facts and figures of the computer industry has shown a healthy growing industry with resilience to economic challenges. We reviewed a broad spectrum ranging from the further extension of Moore as well as the plethora of options for beyond Moore through newly emerging technologies and new computing paradigms. The chart below shows a selection of the rich bouquet of potential present and future paradigms.

COMPUTING PARADIGMS
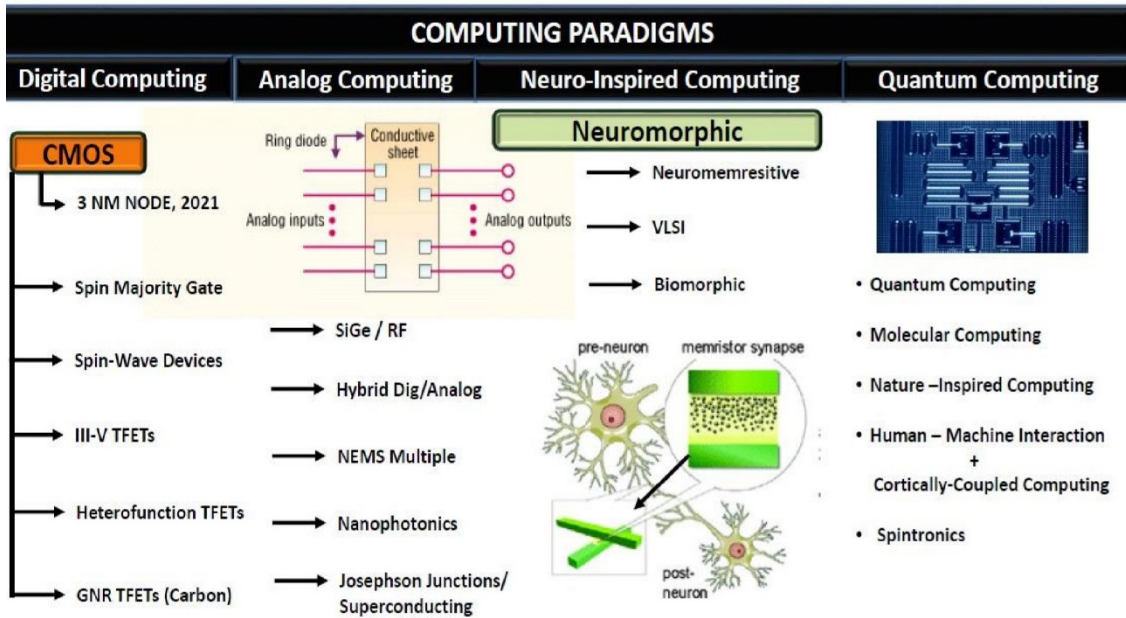
The impact of the pandemic crises shattered previous basic assumptions and made all forecasting difficult. Let us assume that in spite of unprecedented political actions, the related unemployment and debt avalanche the ingenuity and commitment of academic and all shareholders will overcome the present problems and exposures and enable the realization of the promising future outlined.

There is no shortage of ideas and potential and this should sum up to rebooting the IT revolution.

7. References

APPLE, Annual Report, 2019

ALPHABET, Annual Report 2019

AMAZON, Annual Report, 2019

BARDON, M .Garcia, IMEC 2019

BAKKERS, E. M. T. et alii, Direct Bandgap Emission from Hexagonal Ge and SiGe Alloys, 2020

BOTTI Group, Silicon Laser: Efficient light emission from direct band hexagonal SiGe Nanowires FSU

CHROUST G., ICT for resilience of Systems IDIMT 2015

CNRS Strasbourg and Marseille, Nature Communications 2019

MOORE S.K., Ferroelectric Semiconductors Could Mix Memory and Logic

NAKATSUJI S., Univ. Tokyo, MEINERT M, TU Darmstadt, Antiferromagnets the next step, IEEE Spectrum 2020

NDIA, Trusted Microelectronics, Joint Working Group 2017

NEREID, NanoElectronics Roadmaps for Europe, EU Horizon2020, 2019

REED, Technology and Information Services Inc. 2019

SOLOMON P., Analog Neuromorphic Computing using programmable resistor arrays, Solid State Electronics 2019

WALDROP M., More than Moore Nature 530,144

YE LI G., Ferroelectric Semiconductors Georgia Tech.

ZHURUN et LI, Vortex lasers may be a boon for data, Science, May 2020

DIGITALIZATION AND INDUSTRY 4.0

DIGITALIZATION AND INDUSTRY 4.0: ASPECT OF INFORMATION SECURITY

Luděk Novák

Petr Doucek

Lea Nedomová

Jiří Hološka

Faculty of Informatics and Statistics
University of Economics, Prague
ludekn@email.cz, doucek@vse.cz, nedomova@vse.cz, jiri.holoska@vse.cz

Keywords

Information security, security auditing, ISO/IEC 27001:2013

Abstract

Building security management systems is essential to protect the company's assets from losses. Such protection is essential in the dynamic development determined by digitization in Industry 4.0. In our article we analyze data from 152 information security audits performed in Czech and Slovak companies in 2019 in order to identify the most problematic areas in information security management systems. The data – audit findings – were analyzed based on the size of the audited company and the type of audit. We divided the companies into small (up to 50 employees), medium-sized and large (over 250 employees) companies. We divided the audits into four categories – initial, periodic, certification and others. The audits were performed in compliance with the ČSN EN ISO 19011 standard according to the ISO/IEC 27001: 2013 standard. The data were analyzed in MS Excel using contingency tables. The results show that the "A12" category – Operations Security – is the biggest problem for today's organizations. The "A18" category – Compliance – is another problematic area. The positive conclusion is that the shortcomings identified by the audit are not serious shortcomings that could jeopardize the companies' safe operation.

1. Introduction

The Digital Economy with its features as digitalization and Industry 4.0 has been defined as the worldwide network of economic activities enabled by information and communication technologies (ICT) (Industry, 2016). It can also be defined more simply as an economy based on digital technologies (Lin, Chiang, 2011; Hanclova at al., 2015). The whole development of the regional and global economy is closely connected with the rush penetration of ICT into the world economy and pressures the education system to prepare graduates of all levels for more complex knowledge and skills for their future jobs (Mangir, Erdogan, 2011). Nowadays, it is no longer necessary to educate people in what 4.0 trends are and what they can bring. 4.0 trends are already in full swing in companies, at least in some partial ways, and what's more, their dynamics correspond to the overall higher pace of changes (Kuncova, Sekničkova, 2019). And it is not only about

technological stimuli, but also about other demographic and climate stimuli. Therefore, the question is now how quickly these trends will penetrate into the practical life of businesses and society as a whole (Manďák, Nedomova, 2014; Kuncova, Doucek, Novotný, 2018).

Various preparedness indexes and maturity models can facilitate and speed-up companies' decision-making about where and how fast to build industry 4.0. These indexes and models show not only the companies' actual position but also the position of their competition, both at the macroeconomic and microeconomic levels. The focus thus shifts to tasks related to the implementation of necessary changes and to the specification of not only higher profits but also the main expectations associated with their implementation. For instance, achieving maximum flexibility, increasing the availability of products and services, further cost reductions, lower resources consumption and a lower impact on the environment, etc. (Basl, Doucek, 2018; Basl, Doucek, 2019).

However, the penetration and integration of ICT into economic processes also brings phenomena that their actors and participants consider negative. These include e.g. the areas of shared economy (Svecova, Veber, 2018) that clearly demonstrate that the legal aspects of the economy and businesses significantly lag behind ICT and its implementation in economic processes. Legal shortcomings lie mainly in the area of taxes and liability as well as in the processing of personal data and overall security of processed transactions.

In our article we would like to focus primarily on ensuring the security of data processed by companies. The emphasis on this area is clearly shown by European Union legislation such as Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data General Data Protection Regulation (GDPR, 2016). This regulation came into force on 25 May 2018 and it has been followed up with an amendment to Czech Act No. 101/2000 of Coll., on the protection of personal data, to make sure that the Czech law is in compliance with the EU's law (Novak, Doucek 2017) and with new Czech Act No. 110/2019 of Coll., on the processing of personal data (Act No.110/2019 sb.). In addition to personal data, it is also necessary to protect other classified data that companies process and save in their information systems in order to avoid risks (Eling, Schneell, 2016; Doucek et al., 2020).

With this article, we follow up on a presentation at the IDIMT conference in 2018 and analyze data obtained from independent information security audits performed in Czech and Slovak companies in 2019. Our research questions are as follows:

RQ1: What are the main identified problematic areas (according to the ISO / IEC 27001: 2013 standard) in information security in 2019 based on the size of the audited company?

RQ2: What are the main identified problematic areas (according to the ISO / IEC 27001 standard) in information security in 2019 based on the type of audit?

2. Data Collection and Methodology

In order to find out the answer to these research questions, we had to specify the method of data collection and evaluation in two different methodological areas. The first one was data collection and the second one was the way an audit was performed in different security areas, as specified in the ISO/IEC 27001: 2013 standard.

2.1. Data Collection

The data that we used for our conclusions came from the analysis of information system security audits in various industries in the Czech Republic and the Slovak Republic. Since the number of

employees in the companies differed, we divided the audited companies into three categories according to the number of their employees: small companies with up to 50 employees, medium-sized companies with 50-250 employees and large companies with over 250 employees.

Another criterion was the type of audit performed. The abbreviation IA means initial audit, PA means periodical audit in compliance with the general principles for management systems, RA means recertification audit and Oth means other audits (Purcarea et al., 2011). The last type of audit was another one – Oth. In total, we used 152 different audits. Their numbers are provided in Tab. 1 (by company size) and Tab. 2 (by audit type).

Tab. 1 Data Sample by Company Size

Company size	Number of performed audits
Small	31
Medium	36
Large	85
Total	**152**

Tab. 2 Data Sample by Audit Type

Audit type	Number of performed audits
IA/initial audit	7
PA/periodical audit	78
RA/recertification audit	57
Oth/Other	10
Total	**152**

We also analyzed the audit findings based on the seriousness of identified irregularity. Here again, we used four categories: CAT1 – serious, CAT2 – medium, OBS – observation, OFI – opportunity. This criterion is mentioned only at the end of the article and is not used in our analysis.

2.2. Method of Auditing by Area

Audit definition – for the purposes of this text, an audit means a systematic, independent and documented process to obtain and objectively evaluate evidence from the audit in order to determine the extent to which the audit criteria are met (ČSN EN ISO 19011, 2019) (Kaziliunas, 2008). Audit criteria means a set of policies, procedures and requirements used as a basis against which audit evidence is compared (EN ISO 19011, 2019; Hoy, Solei, 2015).

All security audits were performed in compliance with the provisions of the applicable ISO/IEC 27001: 2013 standard and the identified non-conformities were divided according to individual areas of the ISO/IEC 27001: 2013 (ISO/IEC 27001) standard. The numerically marked audit categories correspond to the sections of this standard are in Tab. 3.

Tab. 3 Numerically Marked Audit Categories – Main Sections

Category	Description
"4"	Context of the Organization – Understanding of the organization, its needs and expectation of interested parties and scope of the information management system.

Category	Description
"5"	Leadership – Leadership and commitment, security policy, organizational roles, responsibilities and planning to achieve them.
"6"	Planning - Action to address risks and opportunities, information security objectives and planning to achieve them.
"7"	Support – Resources, competencies, awareness, communication and documented information.
"8"	Operation – Operational planning and control, information security risk assessment and treatment.
"9"	Performance Evaluation – Monitoring, measurement, analysis and evaluation, internal audit, management review.
"10"	Improvement – Nonconformity and corrective action, continual improvement. (ISO 27001:2013).

The next categories marked with an A before the number correspond to the sections of the annex to this standard are in Tab. 4.

Tab. 4 Numerically Marked Audit Categories – Annex Sections

Category	Description
"A05"	Information Security Policies – To provide management direction and support for information security in accordance with business requirements and relevant laws and regulations.
"A06"	Organization of information security – To establish a management framework to initiate the implementation and operation of information security within the organization and to ensure the security of teleworking and use of mobile devices.
"A07"	Human Resource Security – To assure that employees and contractors understand their responsibilities and are suitable for the roles for which they are considered.
"A08"	Asset Management – Responsibility for Assets, Information Classification, Media Handling.
"A09"	Access Control – Business requirements on access control, User access management, User responsibilities, system and application access control.
"A10"	Cryptography – To ensure proper and effective use of cryptography to protect the confidentiality, authenticity and/or integrity of information.
"A11"	Physical and Environmental Security - Secure areas, Equipment.
"A12"	Operations Security - Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.
"A13"	Communication Security – Network security management and information transfer.
"A14"	System acquisition, development, and maintenance – Security requirements of information systems, security in development and support process and data testing.
"A15"	Supplier relationships – Information security in supplier relationship and supplier service delivery management.
"A16"	Information Security Incident Management - To ensure a consistent and effective approach to the management for information security incidents, including communication n security events and weaknesses.
"A17"	Information security aspects of business continuity management – Information security continuity shall be embedded in the organization´s business continuity management system.
"A18"	Compliance – Compliance with legal and contractual requirements, Information security review. (ISO 27001:2013).

In total, there are 21 categories (sections of the standard and its Annex A). In the following text or in the tables with results, the categories are shown in a shortened version only, i.e. either as a number

only – for the categories from the text of the standard or as A and a number – for the areas listed in Annex A of the standard.

The standard statistical functions of MS Excel were used to evaluate the obtained data (Kuncova, Sekničkova 2018).

3. Results

Based on the above criteria, we evaluated the findings based on the size of the company. It is very interesting that no irregularities were identified for area "5" – Leadership – Leadership and commitment, security policy, organizational roles, responsibilities and planning to achieve them for the entire one-year period and therefore it does not appear in the evaluation of audit findings.

3.1. Evaluation by Company Size

First, we analyzed audit data from audits, i.e. audit irregularities by company size. The numbers of identified irregularities in audits based on the size of the company are shown in Tab. 5 and Tab. 6.

Tab. 5 Audit Findings by Company Size – Main Sections

	4	6	7	8	9	10
1 Small	2	2	2	3	1	1
2 Medium	3	2	3	1	4	
3 Large		10	6	1	6	
Total	5	14	11	5	11	1

Tab. 6 Audit Findings by Company Size – Annex Sections

	A06	A07	A08	A09	A10	A11	A12	A13	A14	A15	A16	A17	A18
1 Small			5	4			12		1	1	1	1	6
2 Medium	4		6	6		1	5	1	4			7	7
3 Large	2	4	8	4	1	10	13	8	1	1	6	7	10
Total	6	4	19	14	1	11	30	9	6	2	7	15	23

Tab. 5 and Tab. 6 shows that the biggest problems in small companies involve the "A12" category – **Operations Security –** Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration., i.e. operations security. Smaller companies also have problems in the following areas:

- Category "A18" – Compliance – Compliance with legal and contractual requirements, Information security review.

- Category "A08" – Asset Management – Responsibility for Assets, Information Classification, Media Handling.

- Category "A09" – Access Control – Business requirements on access control, User access management, User responsibilities, System and application access control.

Medium-sized companies have the biggest problems with:

- Category "A17" – Information security aspects of business continuity management – Information security continuity shall be embedded in the organization´s business continuity management system.

- Category "A18" – Compliance – Compliance with legal and contractual requirements, Information security review.

- Category "A08" – Asset Management – Responsibility for Assets, Information Classification, Media Handling.

- Category "A09" – Access Control – Business requirements on access control, User access management, User responsibilities, System and application access control.

In the case of large companies, the biggest problem is in the "**A12**" category – **Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration – operations security. Other areas with identified problems include:

- Category "6" – Planning – Action to address risks and opportunities, information security objectives and planning to achieve them.

- Category "A11" – Physical and Environmental Security – Secure areas, Equipment.

- Category "A18" – Compliance – Compliance with legal and contractual requirements, Information security review.

When evaluating the problematic areas in terms of the number of findings in 2019, we can see the most frequent irregularities in the "**A12**" category **– Operations Security –** Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration for the entire time period. We can also see a lot of irregularities in the "**A18**" category – Compliance – Compliance with legal and contractual requirements, Information security review and "A08" category – Asset Management – Responsibility for Assets, Information Classification, Media Handling. Although this area is not significant in any category, it has gradually obtained points in audits in all analyzed categories.

3.2. Evaluation by audit type

Another evaluated criterion that we used when analyzing the data identified by audits included findings broken down by audit type.

Tab. 7 Audit Findings by Audit Type – Main Sections

	4	6	7	8	9	10
IA/Initial Audit			1		1	
PA/Periodical Audit	2	5	5	3	7	
RA/Recertification Audit	1	9	4	2	1	1
Oth/Other	2		1		2	
Total	**5**	**14**	**11**	**5**	**11**	**1**

Tab. 8 Audit Findings by Audit Type – Annex Sections

	A06	A07	A08	A09	A10	A11	A12	A13	A14	A15	A16	A17	A18
IA/Initial Audit	2		1									1	1
PA/Periodical Audit	1	2	8	8		4	17	6	4	1	6	5	15
RA/Recertification Audit	3	2	8	6	1	7	12	3	2	1	1	9	5
Oth/Other			2				1						2
Total	**6**	**4**	**19**	**14**	**1**	**11**	**30**	**9**	**6**	**2**	**7**	**15**	**23**

Let's consider the number of performed audits based on the type of audit. We can see that there were seven initial audits during the entire year. Therefore, it makes no sense to evaluate this category because of a very small data sample. The situation is similar in the case of other types of audit. As the analysis shows, these audits focused on specific security areas and this is why auditors found irregularities.

For our analysis it makes sense to evaluate two categories – PA – periodical audit and RA – recertification audit. In the case of periodical audits, the most frequently identified irregularities were in the following areas:

- Category **"A12"** – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.

- Category "A18" – Compliance – Compliance with legal and contractual requirements, Information security review.

The following two areas are also worth noticing:

- Category "A08" – Asset Management – Responsibility for Assets, Information Classification, Media Handling.

- Category "A09" – Access Control – Business requirements on access control, User access management, User responsibilities, System and application access control.

The most frequent irregularities in recertification audits included the following areas:

- Category **"A12"** – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.

Other areas identified by the analysis include:

- Category "6" – Planning – Action to address risks and opportunities, information security objectives and planning to achieve them.

- Category "A17" – Information security aspects of business continuity management – Information security continuity shall be embedded in the organization´s business continuity management system.

- Category "A08" – Asset Management – Responsibility for Assets, Information Classification, Media Handling.

4. Conclusions

Based on the analysis of 152 findings of information security audits in Czech and Slovak companies in 2019, we reached the following answers to our research questions:

RQ1: What are the main identified problematic areas (according to the ISO / IEC 27001: 2013 standard) in information security in 2019 based on the size of the audited company?

Final answer:

Tab. 8 Identified problematic areas in information security by audited company size

Company size	Major Problematic Categories	Other Categories
Small	"A12" – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.	"A18"
Medium	"A17" – Information security aspects of business continuity management – Information security continuity shall be embedded in the organization´s business continuity management system. "A18" – Compliance – Compliance with legal and contractual requirements, Information security review.	"A08", "A09"
Large	"A12" – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.	6, "A11", "A18"

RQ2: What are the main identified problematic areas (according to the ISO / IEC 27001 standard) in information security in 2019 based on the type of audit?

Final answer:

Tab. 9 Identified problematic areas in information security by audit type

Company size	Major Problematic Categories	Other Categories
IA/initial audit	Data sample too small to make a conclusion	
PA/periodical audit	"A12" – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.	"A08", "A09"
RA/recertification audit	"A12" – Operations Security – Operational procedures, responsibility, Protection form malware, Back up, Logging and monitoring, Control of operational software, Technical vulnerability management, Information system audit and consideration.	6, "A17", "A08"
Oth/Other	Data sample too small to make a conclusion	

The good thing is that no irregularities of type CAT1 or CAT2 were identified during the 2019 audits. It means that security management systems showed no major irregularities or showed irregularities of medium gravity that have no impact on the quality and reliability of the information security

management system in the company. OBS – observation, OFI – opportunity is used mainly for continuous improvement and further development of the security management system in companies.

5. Acknowledgement

6. References

Basl, J., & Doucek, P. (2019). A Metamodel for Evaluating Enterprise Readiness in the Context of Industry 4.0. Information, 10(3). DOI: http://dx.doi.org/10.3390/info10030089

Basl, J., & Doucek, P. (2018). Metamodel of Indexes and Maturity Models for Industry 4.0 Readiness in Enterprises. 26th Interdisciplinary Information Management Talks – IDIMT 2018 Strategic Modeling in Management, Economy and Society. Linz: Trauner Verlag Universität, pp. 33–40.

ČSN EN ISO 19011:2019 Směrnice pro auditování systémů managementu. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví. Třídící znak: 01 0330

Doucek, P. , Pavlíček, L., Sedláček, J., & Nedomova, L. (2020). Adaptation of Password Strength Estimators to a Non-English Environment – the Czech Experience. Computers & Security. 95 (2020), DOI:10.1016/j.cose.2020.101757.

Eling, M., & Schneell, W. (2016). What do we know about cyber risk and cyber risk insurance? The Journal of Risk Finance. 17 (5), pp. 474 - 491.

GDPR, (2016). GDPR Key Changes. Available at: http://www.eugdpr.org/key-changes.html, Retrieved: 2020_04_04.

Hanclova, J., Doucek, P., Fischer, J., & Vltavska, K. (2015). Does ICT capital affect economic growth in the EU-15 and EU-12 countries? Journal of Business Economics and Management, 16(2), pp. 387-406. DOI:10.3846/16111699.2012.754375

Hoy, Z., Foley, A. (2015). A structured approach to integrating audits to create organizational efficiencies: ISO 9001 and ISO 27001 audits. Total Quality Management & Business Excellence, 26(5-6), pp. 690–702, DOI: 10.1080/14783363.2013.876181

Industry 4.0 (2016). Building the digital enterprise, Global Industry 4.0 survey, Price Waterhouse Coopers.

ISO/IEC 27001:2013 – Information technology – Security techniques – Information security management system – Requirements. International Organization for Standardization.

Kaziliunas, A. (2008). Problems of Auditing Using Quality Management Systems for Sustainable Development of Organizations. Technological and Economic Development of Economy, 14(1), pp. 64-75. DOI:10.3846/2029-0187.2008.14.64-75

Kuncova, M., Doucek, P., & Novotný, O. (2018). Penetrace ekonomických ICT služeb do ekonomiky – srovnání zemí V4. Scientific Papers of the University of Pardubice [online]. 2018, 44(3), pp. 151–162.

Kuncova, M., & Sekničkova, J. (2019). Electricity Consumption Cost for Households in the Czech Republic Based on the High and Low Tariff Rates Ratio – Optimization Model. 37th International Conference on Mathematical Methods in Economics 2019. České Budějovice: Ekonomická fakulta, Jihočeská univerzita v Českých Budějovicích, pp. 547–553.

Kuncova, M., & Sekničkova, J. (2018). Multicriteria Evaluation of the Czech Regions from the Selected Economic Activity Aspects Point of View. In: Quantitative Methods in Economic, Multiple Criteria Decision Making XIX. Bratislava: Letra Edu, 2018, pp. 193–200.

Lin, T. W., & Chiang Ch. (2011). The Impacts of Country Characteristics upon the Value of Information Technology as Measured by Productive Efficiency. International Journal of Production Economics, 132 (1), 13-33.

Manďák, J. & Nedomova, L. (2014). Measuring Performance of European ICT Sectors Using Output-Oriented DEA Models. 22th Interdisciplinary Information Management Talks – IDIMT 2014 Networking Societies – Cooperation and Conflict. Linz: Trauner Verlag universitat, pp. 79–86.

Mangir, F., & Erdogan, S. (2011). Comparison of Economic Perforamnce among Six Countries in Global Financial Crisis: The Application of Fuzzy TOPSIS Method. Economics, Management and Financial Markets, 6 (2), 122-136.

Novák, L., & Doucek, P. (2017). Regulation of Cyber Security in the Banking Sector. In IDIMT-2017 Digitalization in Management, Society and Economy, Linz: Trauner Verlag Universität, pp. 49-54.

Purcarea A.-A., Tiganoaia, B., & Petrea G. (2011). Considerations Regarding the Implementation and Certification Within an Organization of an Information Security Management System. 5th International Conference of Management and Industrial Engineering (ICMIE 2011). Editura Niculescu, Bucuresti, pp. 106-114.

Svecova, L. & Veber, J. (2018). Novelty Digitalization, Revolution, Transformation RR Innovation? 26th Interdisciplinary Information Management Talks – IDIMT 2018 Strategic Modeling in Management, Economy and Society. Linz: Trauner Verlag Universität, pp. 435–445.

Zákon č. 101/2000 Sb. ze dne 4. dubna 2000 o ochraně osobních údajů a o změně některých zákonů, In: Sbírka zákonů České republiky 2000, částka 32, 1521-1532. Available at: http://aplikace.mvcr.cz/sbirka-zakonu/

SearchResult.aspx?q=101/2000&typeLaw=zakon&what=Cislo_zakona_smlouvy

Zákon č. 110/2019 Sb., Zákon o zpracování osobních údajů. Dostupné z: https://www.zakonyprolidi.cz/cs/2019-110

INFLUENCE OF TECHNOLOGICAL CHANGES ON DIGITAL COMPETENCES IN ORGANISATIONS

Anastasiia Mazurchenko, Martin Zelenka, Kateřina Maršíková

anastasiia.mazurchenko@tul.cz, martin.zelenka@tul.cz, katerina.marsikova@tul.cz

Keywords

Industry 4.0, digital skills, organisational structure, technological change, employee development

Abstract

In recent years, "digital competences" has become a key term in the discussion of what kinds of skills and perception people should have in a knowledge-based society. New types of jobs are emerging and ways of acquiring knowledge are changing, with the result that basic digital skills are becoming fundamental to be productive, learn and participate in a modern society. The main objective of the paper is to provide a literature overview of terms such as digital literacy, digital skills and competences. Based on this, the authors introduce their own complex framework of key areas influenced by digital competences, key skills and technology trends. These assumptions are further confirmed by the results of secondary data of selected research in relation to digital competences. The paper shows trends in digitalization on the data obtained in 2016 from 3,500 employees from more than 100 countries. The findings support a framework created based on the literature review by authors and emphasise that organisations need to understand the need to respond to technology developments and to seek and support digital skills across processes throughout the organisation.

1. Introduction

Digital technologies and information dissemination have received much attention in recent years due to their substantial impact on the labour market and the competency sets needed in the economy and society. Due to the ongoing fourth industrial revolution (so-called Industry 4.0) not only the technological background, but also the sociological and psychological working environment is reshaping to a great extent in response to exponentially progressing digital transformation (Ghobakhloo, 2020). One the one hand, some tasks in the digital environment can be processed by algorithm and performed by software robots, which leads to the minimisation of human resources in the enterprises (Shaba et al., 2019). On the other hand, the lack of appropriate employee skills and competences needed for doing work effectively in an almost fully digitalized organisation is still one of the main obstacles to better performance in the digital economy (Machado et al., 2019).

Previous research has shown that there is no consistent approach to identifying the terminology in the area of skills and competences needed for successful digital transformation of human resources owing to the fact that the terms digital literacy (Karpati, 2011; Baškarada & Koronios, 2017), digital skill (Van Laar et al., 2020) and digital competence (Ferrari, 2012; Oberländer et al., 2020) are often used interchangeably. Consequently, the objective of this paper is to identify the difference between these concepts based on the literature review and to propose our own framework of digital skills in response to the ongoing digital transformation. The authors also examine some aspects of digitalization in relation to companies' changing processes, structure and roles. The authors introduce several selected findings on the influence of Industry 4.0 and digitalization on employees' skills

development. A theoretical literature review as well as secondary data analysis present the topic of digital competencies and digital skills in business practice.

2. Literature review

With such technologies as artificial intelligence, autonomous robots, big data, augmented reality, internet of things and cybersecurity becoming more widely available, digitalization of business in all its aspects is becoming a fundamental component of Industry 4.0 and a driver for continuous employee professional development (Trașcă et al., 2019). As Stockhinger & Teubner (2019) have identified, *digitalization* is a complex phenomenon that includes social, mobile, analytical and cloud technologies. Despite the fact that a digital working environment is not natural for humans, there is evidence that the human brain is able to combine digital and analogue processing and to adapt to it (Sarpeshkar, 1998). These changes call for identification of such essential dimensions of the employees' potential development as digital literacy, digital skill and digital competency.

The term *digital literacy* often occurs in macroeconomic and political studies and presents a generalised interpretation, according to which it includes a set of basic skills required for working with digital media, information processing and retrieval. *Digital literacy* enables utilisation of social networks for knowledge creation and sharing, and it supports a wide range of professional computing skills (Karpati, 2011). Theoretical backgrounds for data literacy differ and cover specific topics such as reading the data, questioning the data, reading between the data, reading beyond the data and using data (Mason et al., 2016; Wolff et al., 2016). *Data literacy* is acknowledged as an enabler of effective data collection and utilisation for research (Wetherill, 2017, p. 1) and implementation of software projects (Qureshi & Abdulkhalaq, 2015). Even though digital literacy specifies skills needed for work with big data, it does not always have to be a part of the digital competencies frameworks (Mason et al., 2016). According to Ferrari (2012), *digital competence* is a set of knowledge, skills, attitudes, abilities and awareness that are required when using information and communication technologies to perform tasks, solve problems, communicate, collaborate, create content efficiently, critically, creatively, autonomously, flexibly and ethically for work, leisure and socialising. It should be mentioned that digital competences are usually seen in the context of performing specific jobs or tasks (Kasparova, 2019). Another research study confirmed that *digital competencies* include abilities like the analysis, interpretation and application of the information in relevant live content (Van Laar et al., 2020). In such conditions, a definition of *digital competence at work* as a set of basic skills and other characteristics that enable people to accomplish their job tasks regarding digital media successfully looks much more acceptable (Oberländer et al., 2020).

It is becoming clear that digital skills may be perceived as constituting parts of a complex concept of digital competencies. Van Laar et al. (2020) indicate that *digital skills* are the basic skills necessary to use the internet and required to comprehend and use online content. In line with the first digital competence framework created by the European Commission in 2013, *digital skills* are grouped into the following areas: information and data literacy, communication and collaboration, digital content creation, safety and problem solving (ECDL Foundation, 2017). Besides direct impacts on the employees' competencies, digitalization is also forcing organisations to reconsider their organisational structures as innovation technology adoption speed is increasing (Gerten et al., 2019). Therefore every member of the top management has to be integrated into the company's restructuring processes so each of them will make their own contribution to the business's digital transformation (Becker et al., 2018). Table 1 summarises key areas identified by the authors in the literature and shows the impact of Industry 4.0 and digitalization on the companies' structures.

Table 1. Changes related to digitalization in the business environment

Area	Changes in response to digitalization
Business model	Flexible, data-driven, based on digital components, support innovations
Enterprise architecture	Lean methods adoption, knowledge transfer between employees, cross department cooperation supported by integrated systems
Organisational structure	Decentralization of work structures, roles are defined around the work
Organisational culture	Knowledge workers, highly skilled employees, data-driven culture governance, agility, perfect understanding of the data
Managerial practices	Decentralized decision-making, self-organisation, KPIs connected to Industry 4.0, supporting digitalization
Business processes	Processes supporting innovation, digitalization and changes in ICT

The complex view presented by the authors of this paper shows that it is important for organisations to realise that this is a complex issue that affects not only the need for technologies and competences, but also translates into all directions of organisations' functioning (see Table 1).

3. Methodology

The results of recent critical literature research studies on the topic of digital skills (Van Laar et al., 2017; Oberländer et al., 2020) were used in this research paper as a "qualitative basis" for enriching the authors' own literature framework and maximising current knowledge on the topic. The enrichment was used partly as a critical check of the prior research and as a tool for possible broadening of the digital skill framework. The methodology of systematic steps used in this paper is presented in Figure 1.
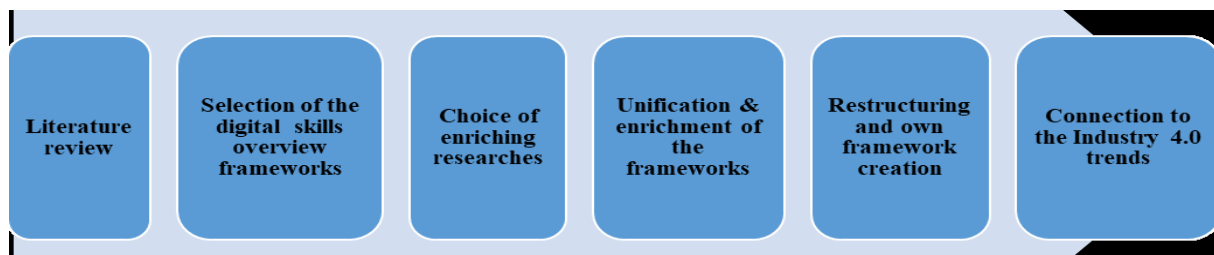
Figure 1. Methodology of the systematic steps used in the paper; Source: Authors' own elaboration

A clear division of skills into hard and soft skills seemed problematic, as Industry 4.0 is characterised by their mutual integration. For this purpose six main areas of the digital skills typical for employees which reflect the nature of the digitalization and technologies connected to it have been chosen by the authors and embedded in the framework (see Table 2). *Digital thinking* is connected to digitalization, using information in appropriate ways and skills needed for programming. *Human-robot interaction* covers skills needed to effectively use software, hardware and also cooperate with robots themselves. *Effective communication* includes efficient sharing of data and using new technologies to deal with colleagues all over the world. *Creative problem solving* combines innovative skill, critical thinking and also creativity during selection of ICT means to solve a given task. *Adaptability* is a response to a continuously changing environment which is demanding on

employees in term of openness and discipline. *Ethical awareness* is a responsible behaviour, maintaining moral standards and security.

Table 2. Framework of digital skill and technology trends in organisations

Area	Typical skills	Technology trends
Digital or algorithmic thinking	Queries, effective search for information, data analysis	Additive manufacturing, big data, artificial intelligence, analytics
Human-robot interaction	Human-robot cooperation, handling of software and applications, ICT-health balance	Robotic process automation, internet of things, additive manufacturing
Effective communication	Multimedia content management, collaboration, cultural aspects, data sharing, explanation of one's own thinking and educating others, networking	System integration, real time response, cloud
Creative problem solving	Innovative capability and creativity, critical thinking, effective selection and ICT usage	Augmented reality, system integration, simulation
Adaptability	Recognising one's own knowledge gaps, goals creation, self-motivation, openness, autonomy, independence, lifelong learning	Robotic process automation, gamification
Ethical awareness	Ethics and morals, responsibility, security and law, netiquette	Cybersecurity, cloud, zero- waste technologies

Source: own elaboration based on (Van Laar et al., 2017; Oberländer et al., 2020; Kasparova. 2019; Colombo et al., 2019; Chetty et. al., 2018; Cahen & Borini, 2020)

Table 2 shows the enhanced framework of digital skills identified by the authors based on the literature review.

4. Results

Data used to confirm key findings from the literature search and for the deeper analysis in this paper originate from the Digital Business Global Executive Study and Research Project conducted by MIT Sloan Management Review and Deloitte in 2016. The poll has been administered to 3,500 business executives, managers, analysts and individual employees in 117 countries across 29 industries. Based on this evidence, some impacts of the digital technologies on the development of employees' skills and organisations' growth can be identified. As Figure 2 shows it is expected that artificial intelligence (18 %) and internet of things (15 %) will be among the most important tools for achieving organisations' digital maturity in the next 5 years from those identified by the authors in Table 2, however the importance of analytics will decrease slightly. Also robotic process automation is expected to grow but not so significantly.
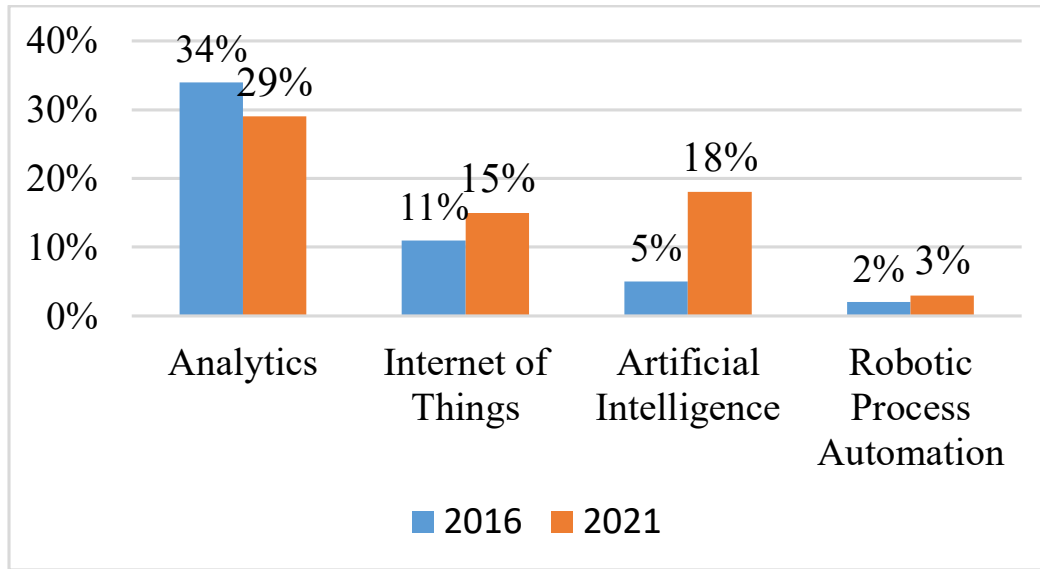
Figure 2. Comparison of selected technologies' importance in years 2016 and 2021; Source: own elaboration based on (Kane et al., 2017)

On the other hand it is interesting that the survey results suggest that social media and mobile technology will become less relevant in adapting the organisations to a changing business environment (9 % and 3 % less respectively). According to this study around half of the respondents (44.9 %) believe that their organisation provides them with the necessary resources and opportunities to develop the skills needed to thrive in a digital business environment. Table 3 introduces the apparent differences between the executives and regular staff in relation to the primary focus of their professional learning.

Table 3. Evaluation of skills development at the workplace by workers

Occupation	Satisfied	Neither satisfied nor unsatisfied	Unsatisfied	Without a clear position
Executives	50.8 %	25.8 %	22.1 %	1.3 %
Employees	40.0 %	25.9 %	32.1 %	2.0 %

Source: own elaboration based on (Kane et al., 2017)

Selected findings from this research confirm that effective utilising of workers' digital knowledge, skills, interest and experience differs to a large extent depending on the organisation's primary industry. This will seem quite clear if we compare results obtained from the companies represented in education and information technologies, and the sector of professional services (see Figure 3). Applying digital skills and knowledge held by employees tends to be more successful in the IT industry whereas educational organisations do not fully use the potential of their staff in this field.
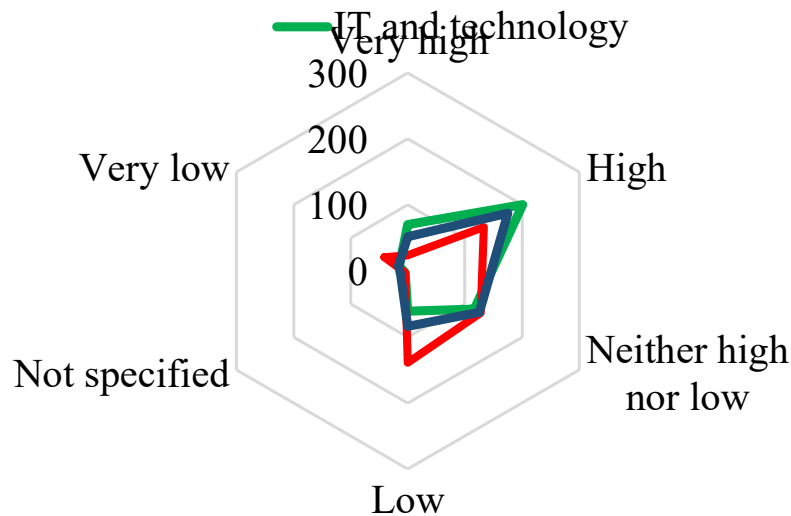
Figure 3. Efficiency of using employees' digital skills and knowledge in selected industries; Source: own elaboration based on (Kane et al., 2017)

The research also shed light on which employees' skills are required for enhancing the company's competitive advantage in a digital business. They particularly include technical (26 %), analytics (19 %) and business skills (10 %). Meanwhile the importance of orientation on changes (6 %), creativity and innovativeness (6 %), critical thinking and problem-solving (1 %) is underestimated.

5. **Discussion and Conclusion**

The framework introduced as well as results presented in the paper based on the literature review provide compelling evidence that jobs have been changing because of the widespread expansion of new digital technologies. This paper also indicates that digital skills are important not only for accessing the labour market, but also for reaping the benefits of the ongoing digital transformation. According to the data by PWC and the Global Manufacturing & Industrialisation Summit (Geissbauer et al., 2016), successful implementation of Industry 4.0 is first of all limited by a lack of digital culture and training needed at the workplace. As the latest survey conducted by the Swinburne University of Technology (Gallagher, 2019) has identified, digital and technology skills are among the most important capabilities for the future of work. Three-quarters of the survey participants feel motivated enough to learn new skills during the next year, because a major part of them (56 %) believe that digital and technology skills will be in the highest demand in the next five years. However, Novak et al. (2018, p. 26) in his study of Central and Eastern Europe's potential for accelerated growth of the digital economy emphasises that employees with advanced technological skills will still be in short supply. The insufficient level of digital skills among the population and the slow pace of their improvement remain a topical problem. Moreover, the digitalization of business is not progressing as fast as it could (European Commission, 2019, p. 7). Overall, the findings from this paper show that technologies tend to have a meaningful impact on the main types of skills essential for the workforce to deal with the challenges in the digital working environment.

The paper aimed to bring an overview of the authors' research and primary data collection in Czech companies. Some limitations of this preliminary study are worth noting. The paper used data from the Digital Business Global Executive Study gathered mainly from organisations with at least a 10-year history (78 %), more than 500 employees (56 %) and also respondents employed in American companies (32 %). Future work should therefore include realisation of the authors' own research

built on the current findings intended to evaluate the actual state of employees' digital skills and demand on employees' competences due to the digitalization in Czech enterprises.

6. Acknowledgements

7. References

Agostini, L., & Filippini, R. (2019). Organizational and Managerial Challenges in the Path Towards Industry 4.0. European Journal of Innovation Management, 22(3), 406-421. DOI: 10.1108/EJIM-02-2018-0030

Baškarada, S., & Koronios, A. (2017). Unicorn Data Scientist: The Rarest of Breeds. Program electronic library and information systems, 51(1), 65-74. DOI: 10.1108/PROG-07-2016-0053

Becker, W., Schmid, O., & Botzkowski, T. (2018). Role of CDOs in the Digital Transformation of SMEs and LSEs-An Empirical Analysis. In Proceedings of the 51st Hawaii International Conference on System Sciences 2018. (pp. 4534-4543). Red Hook, USA: Curran Associates. DOI: 10.24251/HICSS.2018.573

Cahen, F., & Borini, F. M. (2020). International Digital Competence. Journal of International Management, 26(1), 1-17. DOI: 10.1016/j.intman.2019.100691

Cenamor, J., Rönnberg Sjödin, D., & Parida, V. (2017). Adopting a Platform Approach in Servitization: Leveraging the Value of Digitalization. International Journal of Production Economics, 192, 54–65. DOI: 10.1016/j.ijpe.2016.12.033

Chetty, K., Qigui, L., Gcora, N., Josie, J., Wenwei, L., & Fang, C. (2018). Bridging the Digital Divide: Measuring Digital Literacy. Economics, 12, 1-21. DOI: 10.5018/economics-ejournal.ja.2018-23

Colombo, E., Mercorio, F., & Mezzanzanica, M. (2019). AI Meets Labor Market: Exploring the Link Between Automation and Skills. Information Economics and Policy, 47, 27–37. DOI: 10.1016/j.infoecopol.2019.05.003

ECDL Foundation (2017). ECDL and DigComp: Describing, Developing and Certifying Digital Competence. Retrieved from http://ecdl.org/media/digcomp_brochure.pdf

European Commission. (2019). Index Digitální Ekonomiky a Společností (DESI). Zpráva z roku 2019. Czech Republic. Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59981

Ferrari, A. (2012). Digital Competence in Practice: An Analysis of Frameworks. JRC Technical Reports. Luxembourg: Publication Office of the European Union. ISBN 978-92-79-25093-4. DOI: 10.2791/82116

Gallagher, S. (2019). Peak Human Potential: Preparing Australia's Workforce for the Digital Future. Hawthorn, Australia: Centre for the New Workforce Swinburne University of Technology. Retrieved from https://www.swinburne.edu.au/media/swinburneeduau/centre-for-the-new-workforce/cnew-national-survey-report.pdf

Geissbauer, R., Arora, J., & Khurana, A. (2016). Industry 4.0: Building the Digital Industrial Enterprise. PWC and Global Manufacturing & Industrialisation Summit. Retrieved from https://www.pwc.com/m1/en/publications/documents/middle-east-industry-4-0-survey.pdf

Gerten, E., Beckmann, M., & Bellmann, L. (2019). Controlling Working Crowds: The Impact of Digitalization on Worker Autonomy and Monitoring Across Hierarchical Levels. Jahrbücher für Nationalökonomie und Statistik, 239(3), 441–481. DOI: 10.1515/jbnst-2017-0154

Ghobakhloo, M. (2020). Industry 4.0, Digitization, and Opportunities for Sustainability. Journal of Cleaner Production, 252(119869), 1-21. DOI: 10.1016/j.jclepro.2019.119869

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2017). Achieving Digital Maturity: Adapting your company to a changing world. MIT Sloan Management Review and Deloitte University Press. Retrieved from https://sloanreview.mit.edu/2017-digital-business-interactive-tool/guide/

Karpati, A. (2011). Digital Literacy in Education. Moscow: UNESCO Institute for Information Technologies in Education. Retrieved from https://iite.unesco.org/pics/publications/en/files/3214688.pdf

Kasparova, E. (2019). Digital Skills and Their Development in the Czech Republic. Periodicals of Engineering and Natural Sciences, 7(2), 637–643.

Machado, C. G., Winroth, M., Carlsson, D., Almström, P., Centerholt, V., & Hallin, M. (2019). Industry 4.0 Readiness in Manufacturing Companies: Challenges and Enablers Towards Increased Digitalization. In P. Butala, E. Govekar and R. Vrabič (Eds.), Proceedings of the 52nd CIRP Conference on Manufacturing Systems (CMS) 2019. (1113–1118). Amsterdam, Netherlands: Elsevier Procedia. DOI: 10.1016/j.procir.2019.03.262

Mason, J., Khan, K., & Smith, S. (2016). Literate, Numerate, Discriminate—Realigning 21st Century Skills. In W. Chen, J.-C. Yang, Murthy, S., S. L. Wong and S. Iyer (Eds.), Proceedings of the 24th International Conference on Computers in Education. (pp. 609–614). Taoyuan, Taiwan: Asia-Pacific Society for Computers in Education.

Novak, J., Purta, M., Marciniak, T., Ignatowicz, K., Rozenbaum, K., Yearwood, K., Svoboda, D., Karakolev, T. & Skalsky, M. (2018). The Rise of Digital Challengers. How Digitization Can Become the Next Growth Engine for Central and Eastern Europe. Perspective on the Czech Republic. McKinsey & Company. Retrieved from https://digitalchallengers.mckinsey.com/files/The-rise-of-Digital-Challengers_Perspective-on-CZ.pdf

Oberländer, M., Beinicke, A., & Bipp, T. (2020). Digital Competencies: A Review of the Literature and Applications in the Workplace. Computers & Education, 146(103752), 1-13. DOI: 10.1016/j.compedu.2019.103752

Qureshi, M. R. J., & Abdulkhalaq, A. M. (2015). Increasing ERP Implementation Success Ratio by Focusing on Data Quality & User Participation. International Journal of Information Engineering and Electronic Business, 3, 20–25. DOI: 10.5815/ijieeb.2015.03.03

Ravichandran, T. (2018). Exploring the Relationships Between IT Competence, Innovation Capacity and Organizational Agility. Journal of Strategic Information Systems, 27(1), 22–42. DOI: 10.1016/j.jsis.2017.07.002

Sarpeshkar, R. (1998). Analog Versus Digital: Extrapolating from Electronics to Neurobiology. Neural Computation, 10(7), 1601–1638. DOI: 10.1162/089976698300017052

Schwer, K., & Hitz, C. (2018). Designing Organizational Structure in the Age of Digitization. Journal of Eastern European and Central Asian Research, 5(1), 1–11. DOI: 10.15549/jeecar.v5i1.213

Shaba, E., Guerci, M., Gilardi, S., & Bartezzaghi, E. (2019). Industry 4.0 Technologies and Organizational Design-Evidence from 15 Italian Cases. Study Organizzativi, 1, 9-38. DOI: 10.3280/SO2019-001001

Singh, J., Flaherty, K., Sohi, R. S., Deeter-Schmelz, D., Habel, J., Le Meunier-FitzHugh, K., Malshe, A., Mullins, R., & Onyemah, V. (2019). Sales Profession and Professionals in the Age of Digitization and Artificial Intelligence Technologies: Concepts, Priorities, and Questions. Journal of Personal Selling & Sales Management, 39(1), 2–22. DOI: 10.1080/08853134.2018.1557525

Stockhinger, J., & Teubner, A. (2019). How Digitalization Drives the IT/IS Strategy Agenda. ERCIS Working Paper 31, University of Münster, European Research Center for Information Systems (ERCIS)

Teubner, R. A., & Ehnes, D. (2018). The Corporate IT/IS Function: Competences and Organization for a (Digital) Future. In P. Drews, B. Funk, P. Niemeyer and L. Xie (Eds.), Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI) 2018. (pp. 1825–1836). Lüneburg, Germany: Leuphana Universität Lüneburg.

Trașcă, D. L., Ștefan, G. M., Sahlian, D. N., Hoinaru, R., & Șerban-Oprescu, G.-L. (2019). Digitalization and Business Activity. The Struggle to Catch Up in CEE Countries. Sustainability, 11(8), 1-17. DOI: 10.3390/su11082204

Van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & De Haan, J. (2020). Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review. SAGE Journals, 10(1), 1-14. DOI: 10.1177/2158244019900176

Wetherill, O., Lee, C.-W., & Dietz, V. (2017). Root Causes of Poor Immunisation Data Quality and Proven Interventions: A Systematic Literature Review. Annals of Infectious Disease and Epidemiology, 2(1), 1-7.

Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. The Journal of Community Informatics, 12(3), 9-26.

PERSPECTIVE OF HUMAN RESOURCES ALLOCATION IN CONNECTION WITH INDUSTRY 4.0

Hana Mohelská, Majid Ziaei Nafchi

Faculty of Informatics and Management
University of Hradec Králové;

Faculty of Informatics and Statistics
University of Economics, Prague.

hana.mohelska@uhk.cz; majid.nafchi@vse.cz

Keywords

Industry 4.0, human resources, digitization, labour market

Abstract

The onset of the fourth industrial revolution (Industry 4.0, hereinafter I 4.0) brings a completely new approach to production, facilitates the improvement of key industrial processes and thus affects the growth of labor productivity and competitiveness. At the same time, I 4.0 necessitates changes in the organization of work and affects the lives of employees. The paper aims to create a model for predicting the allocation of human resources in the sectors of the national economy of the Czech Republic in connection with I 4.0. The model visualizes the movement of labor in the economic sectors of the Czech Republic from 2013 to the near future. The main contribution of the contribution is to show the development of employment in the high-tech services sector, which will have an increasing tendency.

1. Introduction

The essence of the fourth Industrial revolution is considered and known as Industry 4.0 (hereinafter I 4), and manufacturing sector has been using it lately by utilizing cyber-physical systems (CPS) in order to grasp high levels of automation (Ziaei Nafchi & Mohelská, 2018). The Cyber-Physical System (CPS) is the basis for smart factories and it makes it possible to interconnect sensors, machines and IT systems within the value chain throughout the boundaries of the enterprise (Kopp & Basl, 2017).

Many studies have been conducted about Industry 4.0 and also about readiness and maturity models. The available studies are mostly from the technological point of view. However, the perspective and predictions of allocation of human resource management in the context of I 4.0 is of some importance, which requires more attention and creates a gap for possible research.

According to Klaus Schwab (2018), who is the chairman and founder of the World Economic Forum, four main impacts can be described in the context of I 4.0. It is about:

1. Customer expectations,

2. Product improvement,

3. Innovation in cooperation,

4. Organizational forms.

Everything is cantered around one point, which is the customers and the subsequent care for them, which needs to be constantly improved. As for the products, they can be increasingly enriched with digital capabilities that will increase their value. New technologies increase their resilience, new innovated data their overall value and the way they are maintained. As far as customers are concerned, it is a service-based world, data through analytical services requires special forms of collaboration, one that takes into account speed and innovation (Benhamou, 2018).

According to a comparison of the World Economic Forum in the survey of potential yield from I 4.0, Czech Republic is ranked 6th overall, which can be considered a significant success. This makes it one of the 25 leading countries in the field of robotics. 100 countries and economies were included in the evaluation. The evaluation reflects not only current preparedness, but also how the trend is evolving in the country (Weltwirtschaftsforum & A. T. Kearney Incorporated, 2018).

The Czech Republic achieved this assessment especially for the high share of the developed automotive industry. Czech companies are actively involved in modern trends, thus ensuring competitiveness. For these reasons, I 4.0 represents a huge potential for the Czech Republic.

According to a study (BSG, 2015), which looked at estimates of the impact of I 4.0 on the manufacturing industry in Germany, employment in this sector will increase quite significantly in the next 10 years. The fastest growth (up to 0.9% per year) is logically expected in engineering, which is a producer of new generation production equipment for the entire production sector. This will create a significant demand for new professions throughout the high-tech sector. These professions will combine mechanical engineering with knowledge of electronics, cybernetics and computer science.

The high-technology sector is frequently mentioned as a typical feature of economic development in a particular territorial unit. Businesses in these industries often use advanced technologies to produce on a large scale, while the development of their products is associated with high research and development (R&D) costs and subsequent innovations. The high-technology sector consists of industries within selected industries in the manufacturing industry (high-technology industry) as well as selected service industries (high-technology services) (CZSO, 2019b).

The OECD has developed a four-way exports classification - high, medium-high, medium-low and low-technology. Classification is based on the importance of research and development expenditures relative to the gross output and value added of different types of industries that produce exported goods. Examples regarding high-technology industries are aircraft, computers, and pharmaceuticals; medium-high-technology includes motor vehicles, electrical equipment and the majority of chemicals; medium-low-technology includes rubber, plastics, basic metals and ship construction; low-technology industries include food processing, textiles, clothing and footwear (OECD, 2019; Mohelska et al., 2020).

In high-tech sectors, most human capital is highly skilled and educated. A large part of the employees in these sectors are scientists, engineers and technicians. From the HR point of view, there will be a major restructuring of work. Trends of the onset of I 4.0 are already emerging with increasing intensity. ICT technologies and digitization are penetrating not only into all economic activities, but also into the everyday life of society. In this context, both the scope and structure of employment are changing and will change significantly, new professions are emerging and will be emerging, other professions will have changes in the requirements for knowledge and skills necessary for their performance, other professions will disappear completely.

2. Theoretical background

2.1. Sectors of the national economy

The sector is the part of the national economy into which the national economy is divided in some respects. Sectors are further divided into public, private, mixed or primary, secondary, tertiary and quaternary sectors (Synek & Kislingerová, 2010).

The economy is traditionally divided into three main sectors:

- the agricultural sector (A), which includes own agriculture, fisheries and forestry,

- the industrial sector (I), which includes the mining and processing industries, the production and distribution of electricity, gas and water, in addition to construction,

- the services sector (S), which includes trade, finance and insurance, transport, storage and communications, education, health, veterinary and social work, and other services.

2.1.1. Primary sector

The primary sector is a field associated with obtaining products or raw materials in a natural way. The primary sector is also referred to as primary production, which includes all branches of human activity that convert natural resources into basic products. Developed economies are characterized by a gradual slowdown in the primary sector. They prefer to buy raw materials to buy or use cheaper labour abroad (Stejskal et al., 2017).

2.1.2. Secondary sector

The secondary sector is associated with the processing of products from the primary sector. It is also referred to as the processing or manufacturing sector, which includes all sectors of human activity that convert raw materials into products or goods. In a developed economy, this sector is rather stagnant. Again, cheap labour is used abroad, where part of the processing of products is being moved or by significant modernization, which gradually reduces staff costs (Stejskal et al., 2017).

2.1.3. Tertiary sector

The tertiary sector includes all services. Services related to economic activities, such as banking, financial management and insurance, are also sometimes mentioned. Developed countries, which have a rich population, have the strongest sector and most people work in it. It generates about 60% of GDP in these countries. The services sector includes virtually all organizations (except agriculture, raw materials processing and manufacturing companies), which subject their business to at least one service sector (Kraftová, 2002).

In the last 20–30 years, there has been a dynamic development in the tertiary sector and two other sectors have begun to be separated, namely the Quaternary sector and the fifth sector. They are also listed in the article for completeness.

2.1.4. Quaternary sector

The Quaternary sector used to belong to the tertiary sector, but in the modern conception of economics it is independent (Kraftová, 2002). The Quaternary sector is also referred to as the knowledge sector, as it is based on high education and requires highly qualified staff. The Quaternary sector is growing mainly in the most developed countries and is a driving force for innovation in other sectors, especially in industry and services. This sector creates new industries, produces

innovative services, products and methods of work. Therefore, it is important for the development of the economy and society as a whole.

2.1.5.   The fifth sector

The fifth sector is often referred to as the development services sector, i.e. services that significantly prompt economic growth with an emphasis on its development. This includes services that preserve and cultivate human potential, i.e. especially science and research.

2.2.   **Labor market in the Czech Republic**

This part maps the labour sectors in the economy, describes the development of the labour market since 2013. That is, since the year when I 4.0 began to appear in the Czech Republic. Subsequently, the structure of employees in terms of economic activity is processed and the shares of employment in certain economic sectors of industry, agriculture and services are shown graphically.

2.2.1.   Structure of employment in industry sectors and according to the classification of occupations in the Czech Republic

Employment in the primary sector (I) increased by 20 thousands in 2018 to 1,483,000 people (Table 1). As in the previous year, this sector accounted for 2.8% of total employment.

In the secondary sector (II - industry incl. Construction) a slight decrease in employment was recorded down to 1,984, 900. The share of this sector in total employment decreased in a year from 38.1% to 37.5%.

Table 1: Structure of employment in the period 2017 - 2018 by sector

Employment in the national economy	2017		2018		Difference 2018-2017 (In thousands)	Index 2018/2017 (%)
	In thousands	%	In thousands	%		
Total	5 221.6	100. 0	5293.8	100.0	72.2	100.0
Primary sector (I)	146.3	2.8	148.3	2.8	2.0	101.4
Secondary sector (II)	1 986.9	38.1	1 984.9	37.5	-1.9	99.9
Of that in construction	391.9	7.5	384.0	7.3	-7.9	98.0
Tertiary sector (III)	3 087.5	59.1	3 160.3	59.7	72.8	102.4

Source: MPSV, 2019; Authors

The decrease was mainly due to a decrease in the number of workers in construction, where after a temporary increase in 2017, employment fell again, by 7,900 workers down to 384,000. The share of people employed in construction decreased by 0.2% to 7.3% of the total employment.

Employment in the Czech Republic decreased mostly in the period 2000 to 2018 in the primary and secondary sectors, specifically in the Agriculture, Forestry and Fishery sectors. The number of employees in these sectors decreased by almost 104,000 and in 2018 only less than 130,000 people

worked in this sector.The huge drop in the number of employees also affected the Textile, Clothing and Leather industry, and the Mining industry halved between 2000 and 2014, when instead of 70,000 in 2000, only less than 36,000 remained employed in 2014.

The growth of employment in industry slowed down significantly, the number of employees increased by only 6,000 up to a total of 1,601,000. Share of industry in total employment decreased by 0.3% to 30.2%. All sections contributed slightly to the increase in industry, most notably manufacturing (an increase of 3,600 to 1,458,400 persons).

The overall growth of employment had a positive effect on the development of employment in the tertiary sector of services (III), where employment increased by 72,800 to 3,160,300 persons.Within this sector however, employment developed differently compared to 2017. The number of employees increased mainly in the section of information and communication activities (by 20,400), in the section of wholesale and retail trade, repair of motor vehicles (by 15,600) and in the section of accommodation and food service activities (by 14,700). Lower employment was recorded only in the section of health and social work (by 5,100) and slightly also in the section transport and storage.

The share of the tertiary sector in total employment reached 59.7%, when it increased by 0.6% compared to 2017. This trend continues but the official data available from public sources for 2019 have not yet been published.

Due to I 4, human resources are shifting from sectors I and II to the sector III, and with the arrival of I 4.0 elements, the number of employees in the high-tech sector that uses these elements is increasing.
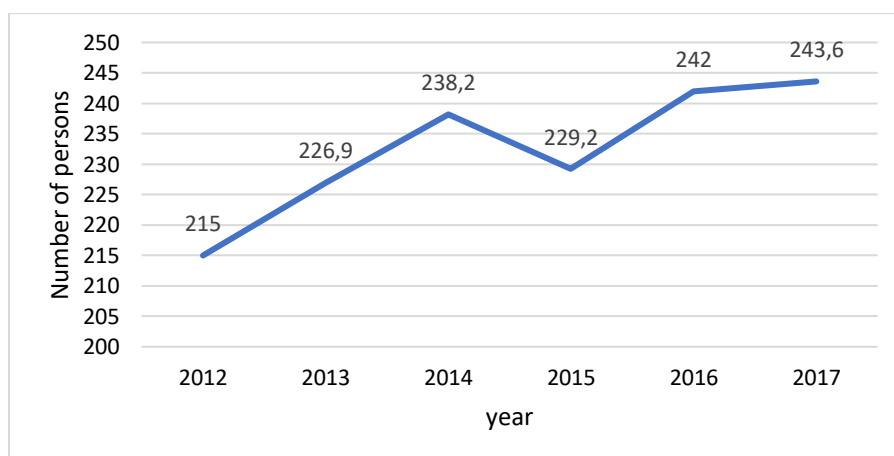
Figure 1: Total employment in the high-tech sector in the Czech Republic (in thousands); Source: CZSO, 2020; Authors.

As shown in Figure 1, it is clear that in 2012 the values are the lowest, as the high-tech sector began to develop only in 2013, when elements of the industrial revolution began to be introduced into business processes and new jobs began to be created.

3. Methodology and Objectives

The aim of the paper is to create a model for predicting the allocation of human resources in the sectors of the national economy of the Czech Republic in connection with I 4.0. The model visualizes the movement of labor in the economic sectors of the Czech Republic from 2013 to the near future.

This model also shows the expansion of robotics from 2013 to the foreseeable future. The limitation of the model is the fact that the development is monitored only in the Czech Republic. All data are taken from publicly available CZSO sources.

The following reservoirs are used in the model (main variables in the model):

- Population - this reservoir visualizes the total population in the Czech Republic set in default values.

- Economically active people - They represent a part of the population or the number of people who are able to work in a group from 15-64 years.

- Robotics - Introduction of a given number of automation machines into the industry.

The model also shows storage facilities in the form of industries (services, industry, and agriculture) and high-tech sectors.

The following flows and quantities are used in the model:

Increase (% population growth in the Czech Republic since 2013), mortality (% population mortality in the Czech Republic since 2013), workers (number of economically active inhabitants), increase in robotics (increasing automation and modernization in sectors), and increase in employees (distribution of able-bodied into individual sectors).

The model was created in the STELLA program. The model itself was visually very extensive, so it is not mentioned in the article. The outputs are shown in the form of graphs, which show the prediction of HR allocation in national economic sectors and the growing number of employees in high-tech sectors of industry and services. The definition and boundaries of the high-tech sector are based on the definition used by CZSO (CZSO, 2019b), which also follows the Eurostat methodology (Eurostat, 2018).

4. Results

The following graphs (Figures 2 and 3) show the predictive situation of the number of workers in economic sectors, without the onset of high-tech industries and the subsequent comparison of the number of workers with the onset of high-tech industries.

The graphs correspond to the expected situation and predict that the number of employees in the national economy will decrease.
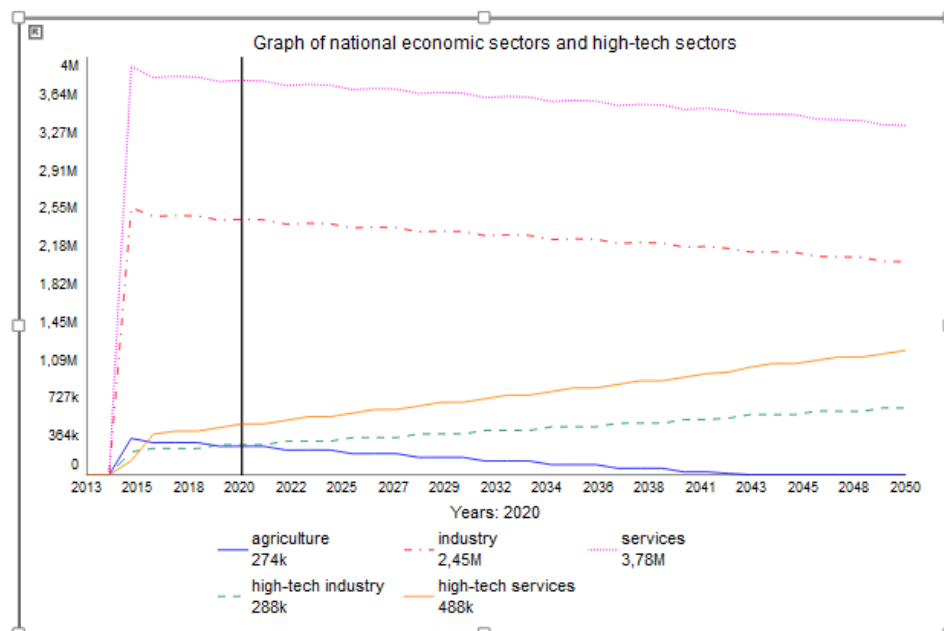
Figure 2: National economic sectors and high-tech sectors (year 2020); Source: Authors.

However, agriculture sector will have the most significant decline. In other national economic sectors, i.e. services and industry, the number of workers will continue to decline in about the same way. In contrast, this declining course shows an increasing number of workers in high-tech sectors, which will eliminate the loss of jobs in existing sectors.
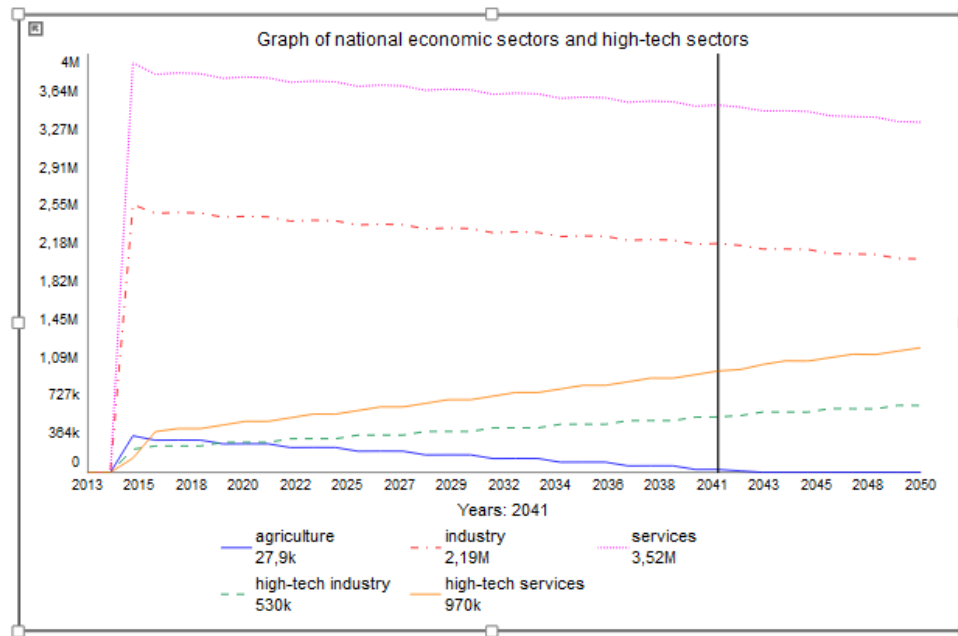
Figure 3: National economic sectors and high-tech sectors (year 2041 - prediction); Source: Authors.

The following figure (Figure 4) shows the economically active people and number of employees since 2013 and subsequent prediction without the onset of high-tech sectors.
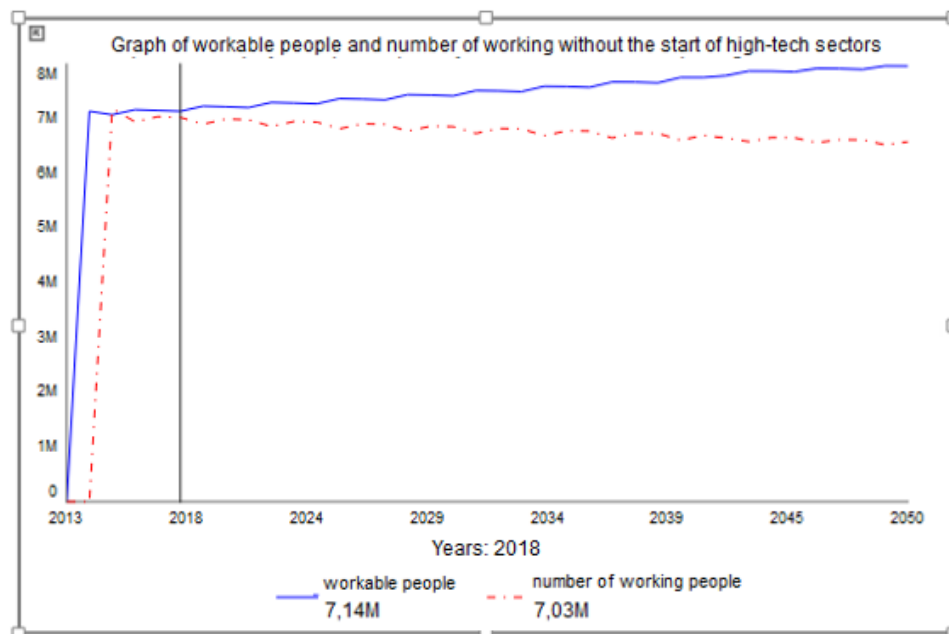
Figure 4: Economically active people and number of working without the start of high-tech sectors (year 2018);

Source: Authors.

Figure 5 shows economically active people and number of employees from 2013 with subsequent prediction until 2050 after setting up high-tech sectors.
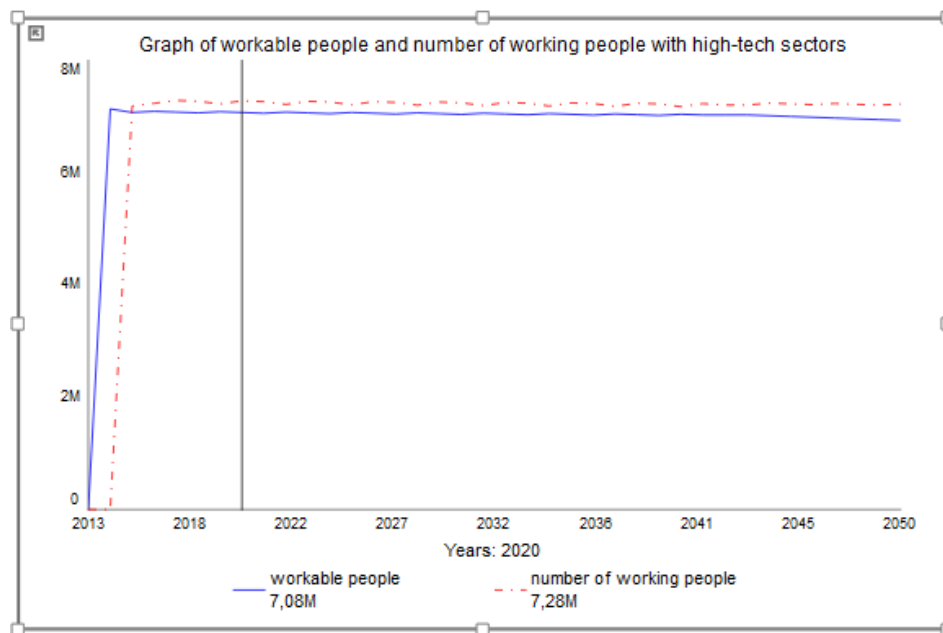
Figure 5: Economically active people and number of working people with high-tech sectors (year 2020); Source: Authors.

5. Discussion and Conclusions

According to Eurostat (2018), the industrial sector in the EU-28 employed 35 million people, representing 15.4% of total employment in 2017. Of this, 2.4 million were employed in the high-tech industry, which is 1.1% of total employment, and the high-tech services sector accounted for three times the number, 3.0% of total employment. The Czech Republic holds on average in the employment of the high-tech sector. Its total employment in this area is 4.6% (services 3.0% and high-tech industry 1.6%). Developed countries such as Ireland (5.4), Finland (4.7%), Estonia (4.5%) and Sweden (4.4%) represented a large difference compared to other countries in terms of services. the lowest values were observed in Greece (2.0%) and Lithuania (2.1%). Only Turkey recorded a value of less than 1%. Ireland had the largest total employment in high-tech sectors, with an overall employment rate of 8.5%. Slovakia appears significantly lower, with a rate of 3.1% of total employment in high-tech sectors.

The issue of future development cannot be accurately predicted or generalized. However, it is necessary to realize that the given problem must always be placed in a specific context and solved with regard to other factors. The basis of successful adaptation to changes in the context of I 4.0 is a proactive approach. The benefit of the contribution is mainly to show the development of employment in the high-tech services sector, which will have an increasing tendency.

6. Acknowledgement

7. References

Benhamou, Salima. The World of Work in 2030: Four Scenarios. In: NEUFEIND, Max, Jacqueline O'REILLY a Florian RANFT. Work in the Digital Age: Challenges of the Fourth Industrial Revolution [online]. London: Rowman & Littlefield International, 2018, s. 259–270 [cit. 2018-07-06]. ISBN 978-1-78660-907-6. Dostupné z: https://policynetwork.org/wp-content/uploads/2018/06/Work-in-the-Digital-Age.pdf

BSG (2015). Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. Retrieved April 17, 2020 from https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx

Bučková, V. (2019). Industry 4.0: A human resource management perspective. University of Hradec Králové, Faculty of Informatics and Management: bachelor thesis.

Czech Statistical Office, (2019a). Employment. Retrieved April 5, 2020, from https://www.czso.cz/csu/czso/13-1132-06-za_rok_2005-3_2__zamestnanost

Czech Statistical Office, (2019b). High-tech sector. Retrieved April 5, 2020 from

https://www.czso.cz/csu/czso/high_tech_sektor?fbclid=IwAR04i98Wa6Zli8mGUOPFqLa9v4YrGjgi7bxsuA15hu9uG4u3nppRQ8L_BI

Czech Statistical Office, (2020). Employment, unemployment. Retrieved April 17, 2020, from https://www.czso.cz/csu/czso/zamestnanost_nezamestnanost_prace

Eurostat (2018). Retrieved from https://ec.europa.eu/eurostat.

Kopp, J., & Basl, J. (2017). Study of the Readiness of Czech Companies to the Industry 4.0. Journal Of Systems Integration (1804-2724), 8(3), 40. http://dx.doi.org/10.20470/jsi.v8i3.313

Kraftová, I. (2002). Finanční analýza municipální firmy. C.H. Beck.

MPSV ČR. (2019). Analýza vývoje zaměstnanosti a nezaměstnanosti v roce 2018 accessed from: Https://www.mpsv.cz/documents/20142/848077/anal2018.pdf/760053b9-9f64-b616-5453-2447ee935b7e. Ministerstvo Práce a Sociálních Věcí ČR, 2019/84292-414/1.

Mohelská, H., Sokolová, M., Cierniak-Emerych, A. & Dziuba, S. T. (2020) Employment in high-technology industries in the european union and job satisfaction - case study Czech Republic and Poland. Economic Research-Ekonomska Istraživanja, 33:1, 639-651, https://doi.org/10.1080/1331677X.2019.1697723

OECD (2019). Retrieved from https://stats.oecd.org.

Schwab, K. (2017). The Fourth Industrial Revolution; Crown Business: New York, NY, USA

Stejskal, J., Kuvíková, H., Mikušová Meričková, B., & Linhartová, V. (2017). Teorie a praxe veřejných služeb. Wolters Kluwer

Synek, M., & Kislingerová, E. (2010). Podniková ekonomika. C.H. Beck.

Weltwirtschaftsforum, & A. T. Kearney Incorporated. (2018). Readiness for the future of production report 2018. https://www.weforum.org/reports/readiness-for-the-future-of-production-report-2018

Ziaei Nafchi, M. & Mohelská, H. (2018). Effects of Industry 4.0 on the Labor Markets of Iran and Japan. Economies, Vol 6, Iss 3, P 39 (2018), (3), 39. doi:10.3390/economies6030039

TO THE RELATION BETWEEN JOB AND STUDY CONTENT: ICT AND ALL STUDENTS

Kristýna Vltavská, Jakub Fischer

Department of Economic Statistics
Faculty of Informatics and Statistics
University of Economics, Prague
kristyna.vltavska@vse.cz, jakub.fischer@vse.cz

Keywords

EUROSTUDENT VII, Job, Logistic regression, Study content

Abstract

Several studies focus on working students. At the EU countries, about half of the students have a paid job during the lecture period. Students have to cover their living costs and would like to gain work experience before graduation. The share of working students differs across countries and their work intensity as well. One of the indicators measured by educational researchers consists of the relation between job and study content, as this relationship has some significant consequences. Our study aims to find the factors influencing work-study relationship for students of the University of Economics, Prague (VSE). We build separate models for ICT students and all students. Using data from the EUROSTUDENT Survey and the logistic regression, the essential factors for VSE students are the age category, work intensity and, especially for all students, the relation of the job before the study to the study content. Unlike other countries, financial difficulties do not influence the work-study association.

1. Introduction

Many students work during their studies across European countries. According to Masevičiuté et al. (2018), 51% of students are engaged in a paid job during the lecture period; two-thirds of them (35% in total) work during the whole lecture period. In the Czech Republic, 49% of students work during the whole lecture period. Engagement in a paid job in the Czech Republic is one of the highest among EU countries. The average weekly time spent on a paid job varies between 14 and 38 hours across countries, less in Western Europe (e. g. 16 hours in Germany and the Netherlands, 17 hours in Switzerland, 23 hours in Austria and France), more in Eastern Europe (Baltic states 33–34 hours, Hungary 35 hours, Poland 36 hours). The Czech Republic lies between these two groups (28 hours per week). For detail, see Masevičiuté et al. (2018), p. 15.

Why half of the Czech students work during the whole lecture period? Although the state and public higher education institutions (hereafter: HEIs) collect no tuition fee, the living costs and other study-related costs increase from year to year. In the Czech Republic, there is no systematic financial support of students (with some small exceptions). Full-time students have 76% of their financial sources from the current job and the support of their families, additional 13% from the previous job (see MŠMT, 2016, p. 88).

For the students of the University of Economics, Prague (hereafter: VSE) the situation is similar. However, their advantage compared to students of e. g. medical study programmes consist of the possibility of having a job during studies. Many of VSE students work for more than 20 hours per week during their studies.

Working during studies brings some issues. Beerkens et al. (2011) point out the low academic standards in Estonia. Unlike in most other countries, many Estonian students from more privileged families more likely to have a full-time job during studies comparing to students from poorer families. According to them, academic standards would change. Hovdhaugen (2015) analyses the impact of employment status on dropout rates. He concludes that students who have a full-time job are less likely to complete their study programme than a student working short part-time or not working at all. The border of the risk of higher dropout lies at 20 hours a week. Tur-Sinai et al. (2017) examine the effect of students' employment on the duration of their studies. For students aged between 22 and 26 years, there is no effect on the extent of employment on the duration of studies, while among the older students, the effect is positive and statistically significant.

According to Masevičiuté et al. (2018), p. 57, almost half of working students (45%) on average have a job related to their study content, from 56% in Estonia to 33% in Ireland. The Czech students are slightly below the European average (43%). At the EU level, master students more often engage in study-related jobs (57%) compared to bachelor students (38%). Work-study relationship increases with age (from 29% for students up to 21 years to 60% for students above 29 years). It differs across study fields (from 57% in education and 56% in ICT to 32% in Natural Sciences). The relation between study and job also depends on the time spent in a paid job (56% for jobs with 20 hours and higher, compared to 37% for shorter part-time jobs), the form of living (38% for living with parents and 48% for not living with them) and the level of financial difficulties (39% for students with financial difficulties and 50% for students without them). According to several studies, students may benefit from having a study-related job (Neill et al., 2004; Tuononen et al., 2015).

This paper aims to find the factors that influence the interconnection between the content of the study program and the current paid job of students in the Czech Republic. We compare students of ICT and all study programmes. The situation of ICT students and graduates is quite specific in the Czech Republic due to the insufficient number of both ICT Specialists and ICT Technicians (Nedomová et al., 2019).

2. Data and Methodology

Data for this analysis comes from the EUROSTUDENT VII survey that was conducted by the Ministry of Education, Youth and Sports (hereafter: Ministry) in spring 2019. The survey was meant for Bachelor and Master students in Czech and foreign language study programs from public, state and private higher education institutions. These students could not be abroad for the mobility or interrupting their studies while the survey took place.

After all data procedures the Ministry got almost 20 thousand questionnaires (CHES, 2019) from which 1,530 were students of VSE. Students from the Faculty of Informatics and Statistics cover 22.6% of VSE students who participated at the survey. We can work only with the VSE dataset as the dataset for all Czech HEIs Has not been published yet. The VSE dataset was weighted according to the structure of students based on the Ministry statistics. Thus, the results can be generalized to the population of all VSE students.

The research question can be analysed by the logistic regression. As the dependent variable (How closely related is/are your paid job(s) to the content of your study program?) has more than two answer categories we choose multinomial logistic regression. We predict the probability of the

occurrence of outcome Y knowing value of independent variables X_{ni}. In the simplest form the equation becomes (Field, 2009):

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_{1i} + \ldots + b_n X_{ni})}} \, ,$$

where $P(Y)$ refers to the probability of Y occurring, e presents the base of natural logarithms, b_0 refers to the constant, $X_{1i}, \ldots, X_{ni}$ present predictor variables and $b_1, \ldots, b_n$ are attached to that predictors. The parameters are estimated using maximum likelihood method.

We use goodness of fit and other tests for the validation of the model (e.g. Hosmer-Lemeshow test, Pseudo R-squared and Wald test). For all formulas see Field (2009).

3. Results

This paper concentrates only on VSE students and its part ICT students at the Faculty of Informatics and Statistics (248 out of 1,530 students). Table 1 shows the structure of students both ICT and all VSE students regarding the main factors we are interested in. ICT students are mainly bachelor students (61.3%), males (73.7%) and up to 25 years (78.3%). More than 70% work during the whole semester (spring 2019) and their job is very closely and closely related to the content of their study program (61.4%). They do not have any financial difficulties at the time of the survey.

On the other hand, VSE students are mainly females (54%). The majority of VSE students are bachelors (57.7%) and up to 25 years (80%). Two thirds of them have job during the whole reference period. For the half of students was their job very closely and closely connected to the content of the study program.

Table 1 represents that not all categories have enough answers thus we join some of them. Regarding the age we joined the last two categories into '25 and more'. Moreover, we joined the answers for the relationship between the content of the study program and the current job into three categories – close (very close and close), average, not at all (last two categories).

We prepared analyses separately for all VSE students (tables 2 to 4) and ICT students (tables 5 to 7). Firstly, we consider all variables from table 1. From our point of view, they mostly influence the decision of students having a job and more concrete having a job which is related to their studies. The dependent variable presents the relationship between paid job and content of current study program with reference category 'Close' as this one has the most answers. All results use 5% level of significance.

Table 1 Descriptive statistics, ICT and all VSE students, %; Source: EUROSTUDENT VII

Variable		ICT	VSE
Study programme	Bachelor	61.3	57.7
	Master	38.7	42.3
Gender	Male	73.7	46.0
	Female	26.3	54.0
Age category	up to 21	33.7	33.8
	22-25	44.6	46.2
	25-30	18.1	17.8
	30 and more	3.5	2.2
Paid job(s) prior to entering higher education for the first time	Yes, I worked contin. for at least 1 year 20h/ w	20.6	14.0
	Yes, I worked for at least one year, less than 20h/week	15.6	19.2
	Yes, I worked, but less than one year	29.8	33.0
	No, I did not work prior to entering higher education	34.0	33.8
How closely related was/were your paid job(s) prior to entering this higher education to your current main study programme?	very closely	12.1	7.3
	2	10.2	10.6
	3	13.2	16.0
	4	17.7	20.5
	not at all	46.8	45.7
Paid job(s) during the current semester	Yes, I work during the whole semester	71.7	66.5
	Yes, I work from time to time during the semester	15.2	15.1
	No, I don't work during the semester.	13.2	18.4
How closely related is/are your paid job(s) to the content of your study programme?	very closely	38.8	25.8
	2	22.6	25.4
	3	11.1	17.3
	4	11.3	14.1
	not at all	16.2	17.4
To what extent are you currently (this semester) experiencing financial difficulties?	very seriously	2.6	2.6
	2	9.5	11.6
	3	23.2	23.2
	4	29.7	23.6
	not at all	35.0	39.0
How well-off financially do you think your parents (or guardians) are compared with other families?	Very well-off	14.6	13.5
	Somewhat well-off	30.4	31.5
	Average	46.7	46.5
	Not very well-off	6.6	7.5
	Not at all well-off	1.8	0.9

3.1. All VSE students

Using likelihood ratio tests, we can state that all variables are needed in the model (table 2). Table 3 shows all Pseudo R-Squared statistics and depicts that the model explains more than 27.2% of the variation in the output (see Nagelkerke R Square). This low value indicates that there are more factors behind the relationship between selected variables and paid job than we have got from the survey. Table 4 presents the final results of the model.

Table 2 Model fitting information, all VSE students

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	427.998			
Final	249.378	178.620	16	.000

Source: Authors' calculation

Table 3 Pseudo R-Squared, all VSE students

Cox and Snell	0.235	Nagelkerke	0.272	McFadden	0.134

Source: Authors' calculation

Table 4 Multinomial logistic regression, 'How closely related is/are your paid job(s) to the content of your study program?' all students

	Average	Not closely
Paid job(s) during the current lecture period		
Yes, I work during the whole lecture period.	0.409***	0.415***
Yes, I work from time to time during the lecture period	-	-
Age category		
Up to 21	2.624**	3.307***
22 to 25 years	1.932**	0.883
25 and more	-	-
How closely related was/were your paid job(s) prior to entering this higher education to your current main study programme?		
Close	0.300***	0.044***
Average	2.067**	0.321***
Not at all	-	-

Source: Authors' calculation; Note: Odds ratios, reference category Not at all; *0.05, **0.01, ***0.001

The results show that we were not able to use all explanatory variables as we expected. The average related job to the content of the study program is mainly influenced by age, paid job during the current semester and the relationship between the job prior entering HEI and its connection to the content of

the current study program. Students between 22 and 25 years old have more likely job in average related (1.932 times) job to their current study program than the one who are older than 25 years. Students younger 21 years have more likely to have a job not at all related to their study program (3.307 times) than older students. Students whose job prior entering the higher education is in average connected to the content of the current study program has more likely (2.067 times) connected their current job and content of the current study program in average than the one with no connection.

3.2. ICT VSE students

For the analysis of ICT students, the explanatory variable 'How closely related was/were your paid job(s) prior to entering this higher education to your current main study programme?' was not statistically significant. Thus, we used only two explanatory variables - age category, do you have (a) paid job(s) during the current semester.

Using likelihood ratio tests we can state that both variables are needed in the model (table 5). Table 6 presents all Pseudo R-Squared statistics and depicts that the model explains more than 21.8% of the variation in the output (see Nagelkerke R Square). Table 7 presents the results of the model.

Table 5 Model fitting information, ICTs students

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	73.090			
Final	37.376	35.714	6	.000

Source: Authors' calculation

Table 6 Pseudo R-Squared

Cox and Snell	0.182	Nagelkerke	0.218	McFadden	0.112

Source: Authors' calculation

Based on the results showed at table 7 we can see that ICT students up to 21 years have more likely not at all connected job and content of their studies (5.003 times) than older ICT students. ICT students who work during the whole semester have less likely average connected job and the content of the current study program (0.268 times) than ICT student who work from time to time.

Table 7 Multinomial logistic regression, 'How closely related is/are your paid job(s) to the content of your study program?' ICT students

	Average	Not closely
Paid job(s) during the current lecture period		
Yes, I work during the whole lecture period.	0.268**	0.195***
Yes, I work from time to time during the lecture period	-	-
Age category		
Up to 21	1.718	5.003***
22 to 25 years	2.881	1.194
25 and more	-	-

Source: Authors' calculation; Note: Odds ratios, reference category Not at all; *0.05, **0.01, ***0.001

4. Discussion

We can compare our results to the situation of the interconnection between work and study fields in other EU countries. For comparison to the EU, we use the EU average, pointed out in the Introduction section. While 35% of EU students have a paid job during the whole semester (Masivičiuté et al., 2018), the share of these students in the Czech Republic is significantly higher (49%), according to the EUROSTUDENT VII results. The percentage of working students with a study-related job in the Czech Republic (43%) is slightly below the EU average (45%). Both at VSE and in the EU countries, the relation between job and study field increases with age. At VSE, age is a more significant factor for ICT students than for all students. Older ICT students at VSE more often have a study-related job.

Furthermore, working during the whole semester leads to closer work-study relation at VSE. It is similar to the EU results when students who work more than 20 hours per week more often have a study-related job. Similarly, to the EU, the share of students with a study-related job is higher for ICT students comparing to students who study different study fields

On the other hand, unlike the EU average, financial difficulties influence work-study relation neither for ICT-students nor for all students. In the EU, the form of living influence having a study-related job. We do not observe this phenomenon for VSE students.

5. Conclusion

The paper aimed to identify the factors influencing the interconnection between job and study field for students of the University of Economics, Prague, having the paid job during their studies. Using the logistic regression, we have built up two separate models for students of ICT study programmes and students of other study fields.

For non-ICT students, the work-study relation depends on the age category, work intensity and relation of paid job before the study. For ICT students, the work-study relation is influenced by the age category and by the work intensity. Age is more significant for ICT students than for the non-ICT ones.

The results are in line with the situation in other EU countries, however concrete factors slightly differ. For VSE students, there is no significant impact of having financial difficulties.

Based on the results of our paper, we recommend to the management of the University of Economics, Prague to actively take into account the fact that the majority of the University students work during their studies. They have several reasons for working during the study: they want to get work experience, cover their financial costs and to solve their financial difficulties. As the first two reasons are essential for many working students, we recommend facilitating the conditions for working within their study field, which more useful for a future position in the job market. According to our results, it means to help students working for the whole semester as the whole-semester working students have a better chance to get a job related to their study field.

For future research, we would like to examine the consequences of the relation between job and study field on the next position on the job market, including the international comparison.

6. References

Beerkens, M., Mägi, E., & Lill, L. (2011). University studies as a side job. Causes and consequences of

massive student employment in Estonia. In: Higher Education, vol. 61, no. 6, pp. 679–692. DOI: 10.1007/s10734-010-9356-0.

The Centre for Higher Education Studies (CHES) (2019), EUROSTUDENT VII first information about the survey (in Czech only), Praha. [Online], Available: https://www.csvs.cz/blog/2019/08/vyzkum-eurostudent-vii-v-ceske-republice-se-dostal-do-faze-zpracovani-dat/ [1 May 2020]

Field, A. (2009). Discovering Statistics Using SPSS. Third edition. Sage, London, 2009. ISBN 978-1-84787-906-6.

Hovdhaugen, E. (2015). Working while studying: the impact of term-time employment on dropout rates. In: Journal of Education and Work, vol. 28, no. 6, pp. 631–651. DOI: 10.1080/13639080.2013.869311.

Masevičiuté, K. et al. (2018). Combining studies and paid jobs. Thematic review. UAB Araneum, Lithuania. ISBN 978-609-468-169-1.

Ministry of Education, Youth and Sports (MŠMT) (2016), EUROSTUDENT VI: Basic results of the survey among students at the Czech HEIs (in Czech only), Praha. [Online], Available: http://www.msmt.cz/uploads/odbor_30/TF/Analyticke_materialy/Eurostudent/E_VI_zaverecna_zprava.pdf. [1 May 2020]

Nedomová, L., Doucek, P., & Maryška, M. (2019). The premises for the development of the digital economy in the Czech Republic. In: IDIMT-2019 Innovation and Transformation in a Digital World, 2019, 41–50. ISBN 978-3-99062-590-3.

Neill, N., Mulholland, G., Ross, V., & Leckey, J. (2004). The influence of part-time work on student placement. Journal of Further and Higher Education, vol. 28, no. 2, pp. 123-137. DOI:10.1080/0309877042000206705.

Tuononen, T., Parpala, A., Mattsson, M., & Lindblom-Ylänne, S. (2015). Work experience in relation to study pace and thesis grade. Investigating the mediating role of student learning. In: Higher Education, vol. 72, no. 1, pp. 41-58. DOI: 10.1007/s10734-015-9937-z.

Tur-Sinai, A., Romanov, D., & Zusssman, N. (2016). The true effect of students' employment on the duration of studies. In: Applied Economics, vol. 49, no. 33, pp. 3328-3340. DOI: 10.1080/00036846.2016.1259749.

INFORMATION TECHNOLOGY CAPABILITY ANALYTICAL RESEARCH EXAMPLE

Alexander Geyda

St. Petersburg Federal Research Center of the Russian Academy of Sciences$

geida@iias.spb.su

Keywords

Abstract

Information technology (IT) capabilities (IC) are defined as the ability to mobilize and deploy IT-based resources in combination with other resources and capabilities. System dynamic capability (DC) is first among capabilities to use in combination with IC. DC is a system's ability to integrate, build, and reconfigure competences to address rapidly changing environments. Here, it is suggested to estimate the indicators of DC and IC based on models and methods to estimate the potential of a system (which has such DC and IC) analytically. In this case DC or IC are characterized by the results of their use for system functioning in changing enthronements. Such results are those parts of a system's potential that are obtained due to given DC and/or IC use. System potential was defined as a system's ability to achieve changing goals in its changing environments. With the use of models and methods built for solving a system's potential problems, it is possible to build functional models in order to estimate a system's potential regarding DC and IC use. As a result of such models and methods, the application of the estimation of DC and IC use indicators becomes possible depending on the parameters and variables of the DC and IC problems to be solved. Such possibility illustrated based on simple example.

1. Introduction

The solutions of many modern practical problems that arise for practitioners require the study of systems' functions, characterized by the necessity to consider changes in these functions because of environmental changes (Aslaksen 2009, 2009). These changes are initiated by informational actions, which, in turn, are caused by environmental influences (Simon 1996, 1996). Such actions are required to implement measurements, change state checks, transfer information, and prescribe further actions, in some cases including actions to change the system itself (Davis et al. 2009, 2009). Information operations are elements of activity whose objectives are to obtain information, not to exchange matter and energy. Information operations are implemented in accordance with certain information technology whose objective is to describe the use of information operations. Unfortunately, the mechanisms of the formation of activity effects and the subsequent formation of operational properties, taking into account the use of information operations, including modern (digital) IT, have not been studied in sufficient detail in order to predict the effects of activity with mathematical models, depending on the selected characteristics of the information operations used, as mathematical problems of evaluation, analysis, and synthesis.

The hypothesis of the research is that the quality of system functioning in changing conditions and with regard of IT use can be evaluated analytically, based on predictive mathematical models. Author

suggested a property of the system potential (or system capability) that describes the effects of the compliance of changing system functioning with changing conditions analytically and predictively. Thus, it describes the performance of the use of information operations to realize functioning in changing conditions. The new property of the CTS potential proposed by author is an operational property that characterizes the CTS's ability to achieve changing (i.e., actual and possible) goals during operation (in a changing environment). Models and method to estimate this properties indicators are illustrated on the base of simplest but complete example of graph-theoretic probabilistic models use (Dechter 2019, 2019) to estimate indicators.

The methodology suggested for the research is analytical estimation of indicators using probabilistic alternative actions networks.

2. Example of IT use and IT capability estimation

Let us consider basis (started according basis plan π_0) network of 2 operations, A_1, A_2, which fulfilled in parallel. The operations can be interrupted due to possible change of the goal by the environment. Goal can be changed from G_0 to G_1 and only once, but information operations required to check for such change. Information operation I_k checks the environment and system states and able to interrupt functioning in order to start new one. Single informational technological operation I_1 to check environment and system states performed. it is fulfilled 1 in parallel to A_1, A_2. Two alternative variants of I_k considered : I_{11}, I_{12} The moment of A_1, A_2 start is equal for both operations, T_0. $T_0 = 0$, deterministic. Basis network illustrated in Figure 1.
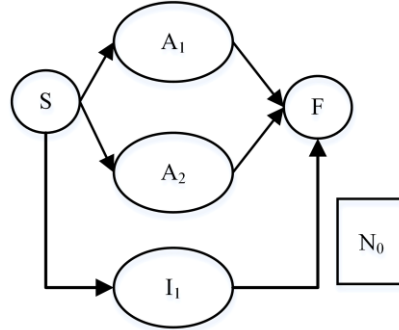
Figure 1: The basis network N_0 of 2 operations

The moment to begin I_1 is T_1; T_1 is deterministic in example; The demands are to finish first operation A_1 by this moment and to not overspend funds C_{01}^d allocated by this moment; Moments T_0 and T_2 are deterministic too; Information operations of special ("dummy") type (start, finish) associated with s_0, f_0. Depending information technology it_k used (it_1, it_2) for information technological operation I_1 (I_{11}, I_{12}) fulfillment (as well as s_j, f_j fulfillment) the system will react in different way on changes in the environment and on its impact on the system. Let us fix basic IT number as it_1 and built model of the system reactions on system environment for it_1.

2.1. Possible sequences of states and technological operations for IT 1 used

Demands at T_2 depends on goal (G_0, G_1) which set by environment to the moment T_1. it_1 used for IT_1 allows to react on changing goal (Figure 2) and to change demands for system functioning. For goal G_0 demands formed by I_1 are to end all operations up to T_2 and not to overspend funds C_{02}^d allocated to this moment; For G_1 demands are to interrupt goal G_0 fulfillment, to start new network of operations and than to finish all technological operations according appropriate network N_1, N_2, N_3, N_4 (depending on state S_1, S_2, S_3, S_4 of the beginning of the goal G_1 achievement there

will be different networks of operations) up to the moment T_2 and to not overspend funds C_{12}^d allocated up to this moment. Funds allocated are supposed independent from state at which system functioning was interrupted due to goal change;
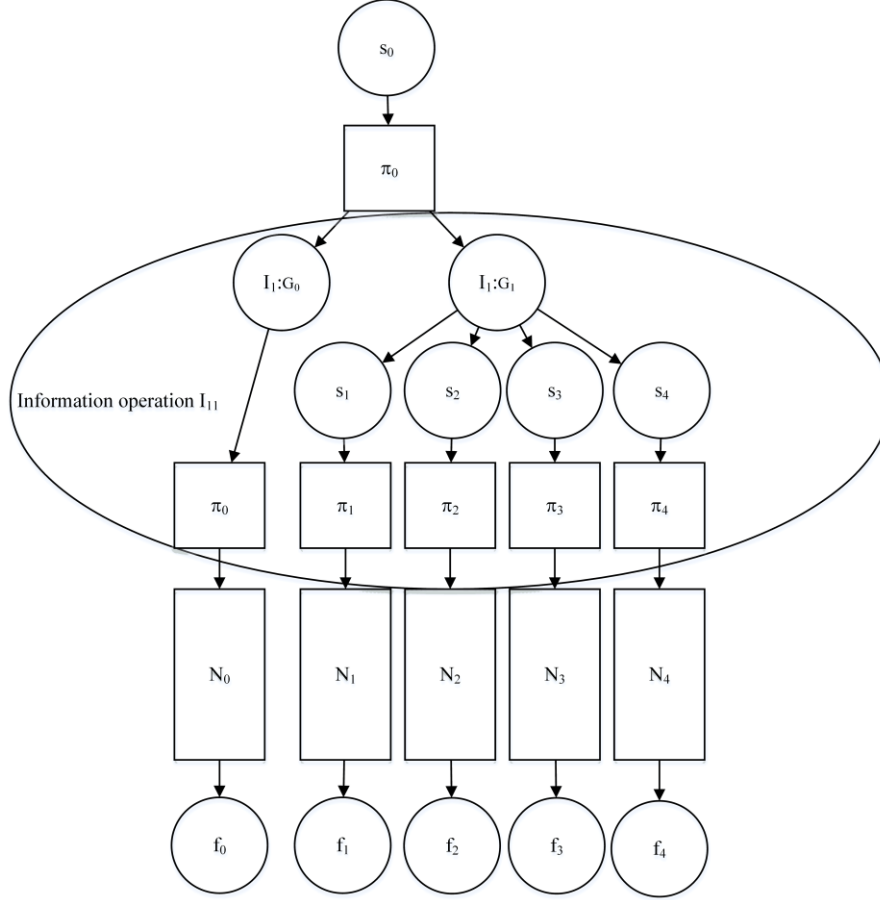
Figure 2: Possible states in a sequences as a result of information operation according it_1 used for I_1

Duration of I_1 (designated as t_1) is negligible compared to A_1, A_2 duration, i.e. $T_1 + t_1 \approx T_1$ and it is supposed that resources are not spent on I_1; At moment T_1 demands are checked according current goal G_0, i.e. according initial goal, which was actual at T_0 moment, before T_1; Probability $P(\hat{A}_1^r)$ of the event $\hat{A}_1^r$ which is - technological operation A_1 realizes (have began but not yet ended) at the moment T_1 is equal to:

$$[P(\hat{A}_1^r(T_1)) = P(\hat{A}_1^b(T_1))(1 - P(\hat{A}_1^e(T_1))) = f_{a1}^b(T_1; a_1^b, b_1^b)(1 - f_{a1}^e(T_1; a_1^e, b_1^e)), \qquad (1)$$

where $f_{a1}^b(T)$ - cumulative probability distribution function of the moment $\hat{T}_{a1}^b$ of the technological operation A_1 beginning at the moment T; Thus,

$$P(\hat{A}_1^b(T_1)) = f_{a1}^b(T_1; a_1^b, b_1^b), \text{where} f_{a1}^b(T_1; a_1^b, b_1^b) - \qquad (2)$$

probability of event $\hat{A}_1^b(T_1)$ that at the moment T_1 technological operation A_1 began.

$$P(\overline{\hat{A}}_1^e(T_1)) = (1 - P(\hat{A}_1^e(T_1))) = (1 - f_{a1}^e(T_1; a_1^e, b_1^e)) - \qquad (3)$$

probability of event $\overline{\hat{A}}_1^e(T_1)$ that at the moment T_1 technological operation A_1 have not yet ended. In this case, probability $P(\hat{A}_2^r)$, of the event $\hat{A}_2^r$, which consist of that technological operation A_2 realizes (i.e. began but not yet ended) at the same moment T_1:

$$P(\hat{A}_2^r) = P(\hat{A}_2^b(T_1))(1 - P(\hat{A}_2^e(T_1))) = f_{a2}^b(T_1; a_2^b, b_2^b)(1 - f_{a2}^e(T_1; a_2^e, b_2^e)), \qquad (4)$$

where $f_{a2}^b(T)$ - cumulative function of probability distribution of the moment $\hat{T}_{a2}^b$ of technological operation A_2 beginning at the moment T;

$$P(\hat{A}_2^b(T_1)) = f_{a2}^b(T_1; a_2^b, b_2^b), \text{ where } \quad f_{a2}^b(T_1; a_2^b, b_2^b), \tag{5}$$

probability of event, that at the moment T_1 technological operation A_2 not began.

$$(1 - P(\hat{A}_2^e(T_1))) = (1 - f_{a2}^e(T_1; a_2^e, b_2^e)) - \text{probability} P(\overline{A}_2^e(T_1)), \tag{6}$$

of that, at moment T_1 technological operation A_2 have not yet ended. There are two possible outcomes of actions of the environment available at T_1 and they are shown as possible results of information operation I_1 according it_0 - information about G_0, G_1 in the form of demands to the effects of functioning by the environment at moments T_2. Probabilities of such outcomes of I_1 are:

$$P(\hat{G}_0 = 0{,}55), P(\hat{G}_1 = 0{,}45). \tag{7}$$

Depending on measured by information operation I_1 the environment and system states there will be different plans for the further operations and different networks of operations realized. Branches are shown at 2 for it_0. Corresponding networks of operations are shown in Figure 3.
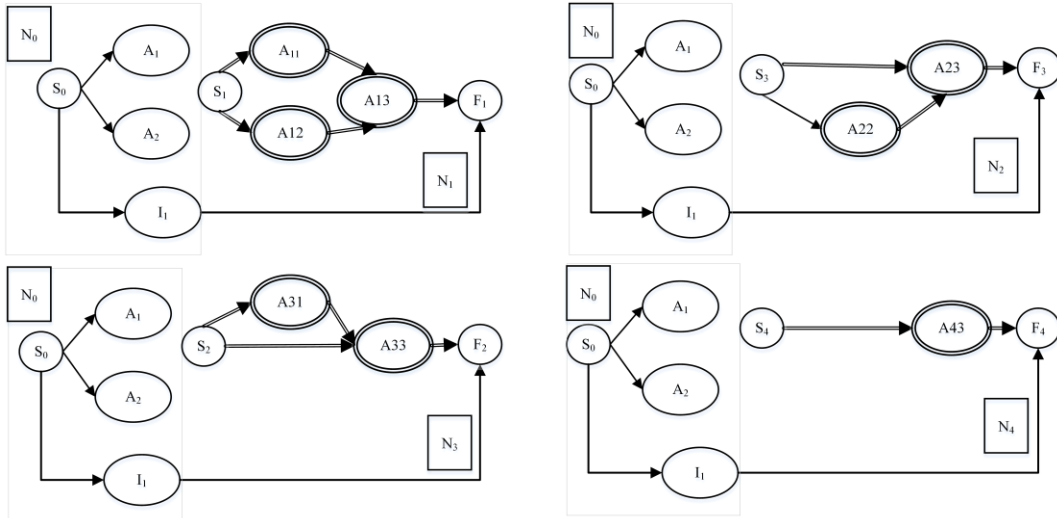
Figure 3: Possible networks as a result of it_1 use to react on the environment change

Probability to realize appropriate branch of events B_0:

$$P(\hat{B}_0) = P(\hat{G}_0)P(\hat{A}_1^r)P(\hat{A}_2^r), \tag{8}$$

i.e. goal not changed (still G_0) at T_1, so demands up to the moment T_2 not changed, that is - up to this moment A_1 and A_2 shall be ended and funds C_{02}^d shall not be overspend. It is supposed for simplicity nothing shall be changed in course of technological operations in this case.

$$P(\hat{B}_1) = P(\hat{G}_1)P(\hat{A}_1^r)P(\hat{A}_2^r), \tag{9}$$

that is, goal changed, first and second operations are fulfilled (started, not ended). B_1 constructed under assumption goal changed from G_0 to G_1 it means, it is needed to interrupt technological operations A_1, A_2, than to restore workplaces W_1, W_2 where A_1, A_2 performed and than, after workplaces restored, to realize technological operation A_{31} of new goal G_1 fulfillment).

$$P(\hat{B}_2) = P(\hat{G}_1)(1 - P(\hat{A}_1^r))P(\hat{A}_2^r), \tag{10}$$

that is, the goal changed, first operation already finished and second operation fulfilled. B_2 constructed under assumption goal changed from G_0 to G_1. It means, it is needed to interrupt

technological operation A_2, than to restore workplace W_2 where A_2 performed and than, after workplaces restored, to realize technological operation A_{32} of new goal G_1 fulfillment).

$$P(\widehat{B}_3) = P(\widehat{G}_1)P(\widehat{A}_1^r)(1 - P(\widehat{A}_2^r)), \tag{11}$$

that is, the goal G_0 changed to G_1 and first operation A_1 realize and second one A_2 - not realized at moment T_1 when goal change and workplaces states checked. It means it is needed to interrupt A_1 and to restore workplace W_1. Next, technological operation A_{33} shall be performed to fulfill goal G_1.

$$P(\widehat{B}_4) = P(\widehat{G}_1)(1 - P(\widehat{A}_1^r))(1 - P(\widehat{A}_2^r)), \tag{12}$$

that is, the goal changed, both operations already finished. It means, nothing to interrupt, it is needed to realize technological operation A_{34} of new goal G_1 fulfillment).

As a result, each branch corresponds to network of technological operations and vector of such networks formed:

$$N := <N_0, N_1, N_2, N_3, N_4>. \tag{13}$$

where N_0-corresponds to case when nothing changed, according B_0 initial plan was fulfilled. Two probabilities where calculated, which correspond to measures of demands and effects compliance for two moments - T_0, T_1. Probabilities was multiplied, than multiplied on B_0 actualization probability. Probability to reach goal G_0 under assumption of branch B_0 realisation computed and multiplied on B_O probabilities. That provided first part of sum to estimate system potential - among other 4 parts.

N_1-goal changed on G_1, according branch B_1 initial plan π_0 realized up to the moment T_1, one probability P_0 as measure of effects and demands correspondence at T_1 counted, than plan π_1 was fulfilled and N_1 realized. Plan was to realize, first, operations A_{11} and A_{12} to restore workplaces W_1, W_2. Because of they fulfilled on different workplaces (W_1, W_2), they was restored in parallel. Next operation A_{13} to fulfill new goal G_1 realized. At the moment T_2 second multiplier of probability to reach the goal G_1 was calculated. Probability to reach goal G_1 under assumption of branch B_1 realisation computed and multiplied on B_1 probabilities. That provided second part of sum to estimate system potential - among other 4 parts.

N_2-goal changed on G_1, according branch B_2 initial plan π_0 realized up to the moment T_1, one probability P_0 as measure of effects and demands correspondence at T_1 counted, than plan π_2 was fulfilled and N_2 realized. Plan was to realize, first, operations A_{22} shall be realized to restore workplace W_2. Workplace W_1 was already in right state at T_1. Next operation A_{23} to fulfill new goal G_1 realized. At the moment T_2 second multiplier of probability to reach the goal G_1 was calculated. Probability to reach goal G_1 under assumption of branch B_2 realisation computed and multiplied on B_2 probabilities. That provided third part of sum to estimate system potential - among other 4 parts.

N_3-goal changed on G_1, according branch B_3 initial plan π_0 realized up to the moment T_1, one probability P_0 as measure of effects and demands correspondence at T_1 counted, than plan π_3 was fulfilled and N_3 realized. Plan was to realize, first, operations A_{13} shall be realized to restore workplace W_1. Workplace W_2 was already in right state at T_1. Next operation A_{33} to fulfill new goal G_1 realized. At the moment T_2 second multiplier of probability to reach the goal G_1 was calculated. Probability to reach goal G_1 under assumption of branch B_3 realisation computed and multiplied on B_3 probabilities. That provided fourth part of sum to estimate system potential - among other 4 parts.

N_4-goal changed on G_1, according branch B_4 initial plan π_0 realized up to the moment T_1, one probability P_0 as measure of effects and demands correspondence at T_1 counted, than plan π_4 was fulfilled and N_4 realized. No need to restore workplaces W_1, W_2, they are both restored. Operation A_{43} to fulfill new goal G_1 realized. At the moment T_2 second multiplier of probability to reach the goal G_1 was calculated. Probability to reach goal G_1 under assumption of branch B_4 realisation

computed and multiplied on B_4 probabilities. That provided fourth part of sum to estimate system potential - among other 4 parts.

As a result matrix obtained, which describes with which probabilities which network fragment will be realized under different conditions of fulfillment and interruption:

$$[(N_0, P(\hat{B}_0), (N_1, P(\hat{B}_1), (N_2, P(\hat{B}_2), (N_3, P(\hat{B}_3), (N_4, P(\hat{B}_4)] \quad (14)$$

Accordingly, networks and demands (according goals G_0, G_1) and next, the measures of the correspondence are calculated. It reflects measure of partial achievements of the goal, of successful transitions to new functioning and further functioning according possible alternatives:

$$
\begin{aligned}
W &= \quad < W_0, W_1, W_2, W_3, W_4 >, \text{where} \\
W_0 &= \quad W_{01} W_{02}; \\
W_1 &= \quad W_{01} W_1; \\
W_2 &= \quad W_{01} W_2; \\
W_3 &= \quad W_{01} W_3; \\
W_4 &= \quad W_{01} W_4.
\end{aligned}
\quad (15)
$$

As well, probabilities of branches (and so - networks realization) can be represented as matrix B according equation (13) used for potential and IT capability estimation.

Than, scalar indicator of system potential in example considered, giving information operation I_{11} used:

$$\phi(I_{11}) = W_0 P(\hat{B}_0, I_{11}) + W_1 P(\hat{B}_1, I_{11}) + W_2 P(\hat{B}_2, I_{11}) + W_3 P(\hat{B}_3, I_{11}) + W_4 P(\hat{B}_4, I_{11}). (16)$$

2.2. Possible sequences of states and technological operations for IT 1 used

Let us consider another IT operation, I_{12} fulfilled in accordance with IT it_2. Its difference is in that, it makes changes to functioning for initial goal G_0, depending on states which was measured by I_{12}. See Figure (4) for details. Depending on the state measured different plans $\pi_{01}, \pi_{02}, \pi_{03}$ will be formed or - alternatively, decision to do nothing (plan, which corresponds to f_3). Next, plans will be fulfilled and appropriate branches of states and actions B_{01}, B_{02}, B_{03} will be realized - or - functioning will stop successfully at T_1 which corresponds to B_{04} and appropriate probability. As a result, scalar indicator of system potential in example considered, giving information operation I_{12} used by analogy with equation (16):

$$\phi(I_{12}) = W_0 P(\hat{B}_0, I_{12}) + W_1 P(\hat{B}_1, I_{11}) + W_2 P(\hat{B}_2, I_{12}) + W_3 P(\hat{B}_3, I_{12}) + W_4 P(\hat{B}_4, I_{12}). (17)$$

Difference between $\phi(I_{12})$ and $\phi(I_{11})$ suggested as IT capability measure:

$$\Phi(I_{12}, I_{11}) := \psi(I_{12}) - \psi(I_{11}). \quad (18)$$
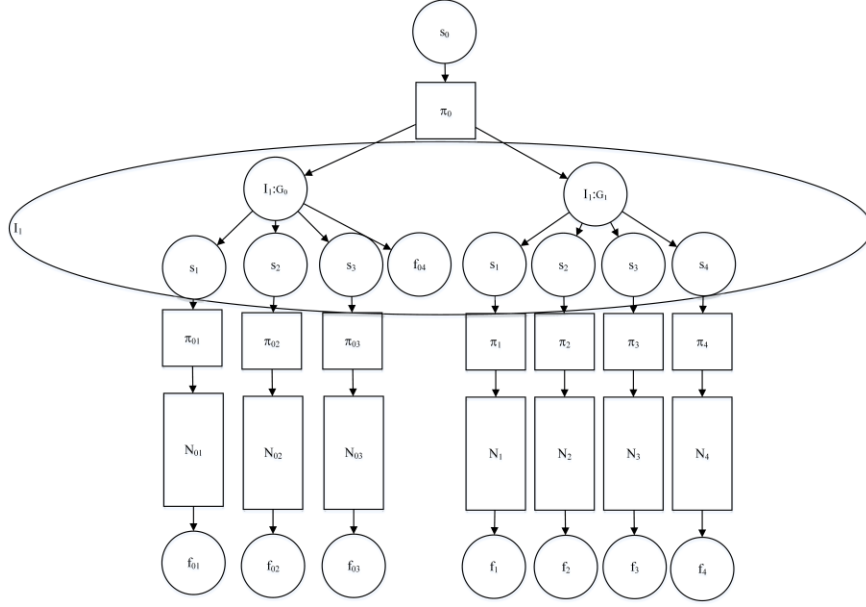
Figure 4: Possible networks as a result of information operation I_1 according it_2 use

3. Research of IT capability based on the system potential

In the general case matrix 15 and array 13 will get the form of multidimensional array $W_{[Q,A,J,U]}$, which allows to get w_{qaju}, p_{qaju} (i.e. W and P) for each branch $B_q^r \in B^r$ and goals sequences G_r $it_k \in I$:

$$
\begin{aligned}
W_{[Q,A,J,U]} &= < w_{qaju}, p_{qaju}, u = \overline{1,U}, j = \overline{0,J}, \\
q &= \overline{1,Q}, a = \overline{1,A} >.
\end{aligned}
\tag{19}
$$

It is possible to use any or few characteristics of such multidimensional discrete random vector or some kind of probabilistic mix as the scalar CTS potential indicator $\psi_1(I_a)$ as the function of IC I_a or DC D_c used. For example:

$$
\psi_1(I_a) = \sum_{q=1}^{Q} \left(\prod_{j=1}^{J} \sum_{u=1}^{U} (w_{qaju}(I_a, S_q) p_{qaju}(I_a, S_q)) \right) p_q;
\tag{20}
$$

$$
\psi_1(D_c) = \sum_{q=1}^{Q} \left(\prod_{j=1}^{J} \sum_{u=1}^{U} (w_{qcju}(D_c, S_q) p_{qcju}(D_c, S_q)) \right) p_q;
\tag{21}
$$

where w_{qaju}, p_{qaju} at (20) are taken from $W_{[Q,A,J,U]}$ at (19) and $w_{qcju}, p_{qcju}, c \in \overline{1,C}$ are taken from appropriate $W_{[Q,C,J,U]}$. Alternatively, the CTS potential indicators $\psi_2(I_a), \psi_2(D_c)$ with regard of IC and DC use can be evaluated as guaranteed values (based on pessimism criteria) :

$$
\psi_2(I_a) = \sum_{q=1}^{Q} \left(\prod_{j=1}^{J} \min_{\forall u \in \overline{1,U}} (w_{qaju}(I_a, S_q)) \right) p_q;
\tag{22}
$$

$$
\psi_2(D_c) = \sum_{q=1}^{Q} \left(\prod_{j=1}^{J} \min_{\forall u \in \overline{1,U}} (w_{qcju}(D_c, S_q)) \right) p_q;
\tag{23}
$$

IT I_a capability indicator $\Phi(I_a, I_1)$ compared to basic - IT I_1 - can be estimated as difference:

$$
\begin{aligned}
\Phi_1(I_a, I_1) &:= \quad \psi_1(I_a) - \psi_1(I_1), \\
\text{or}, \Phi_2(I_a, I_1) &:= \quad \psi_2(I_a) - \psi_2(I_1).
\end{aligned}
\tag{24}
$$

DC D_c indicator $\Phi(D_c, D_1)$ compared to basic - for example, not dynamic (zero-level) capability D_1 - can be estimated as difference:

$$\Phi_1(D_c, D_1) := \quad \psi_1(D_c) - \psi_1(D_1),$$
$$\text{or, } \Phi_2(D_c, D_1) := \quad \psi_2(D_c) - \psi_2(D_1). \tag{25}$$

Equations 19-25 can be used, similarly, to estimate indicators of the competitive performance, organizational capabilities. The strategic alignment of IC and DC can be estimated as the difference of aligned DC, IC CTS potential and the CTS potential for not aligned DC, IC. Let us consider typical practical problem of capability research. Given known historical data D about the CTS usage routines in the past and possible routines of IT capability from IC, DC $DC \in DC$ it is needed to find out which optimal capabilities $C_{dca}^{Opt} = D_d \cup D_c \cup I_a$, where $D_d \in D, I_a \in I, I \in IC, D_c \in DC$ shall be used and what is the best plan $\pi^{Opt}(C_{dca}) \in \Pi_{dca}$ to use to align their usage in the changing environment:

$$
\begin{aligned}
Given: \quad & C_{dca} \in C; C = D \cup IC \cup DC, \Pi_{dca} \in P; \\
Find: \quad & C_{dca}^{Opt}, pi^{Opt}(C_{dca}^{Opt}): \\
C_{dca}^{Opt} \quad & \in \underset{C_{dca} \in C}{ArgMax} \Phi(D_c, I_a, D_0, I_a, \pi_1); \\
\pi^{Opt}(C_{dca}^{opt}) \quad & \in \underset{\pi_{jdca} \in \Pi_{dca}}{ArgMax} \Phi(\pi_{jdca}, C_{dca}^{Opt}); \\
Where: \quad & \Phi \in \{\Phi_1, \Phi_2\}.
\end{aligned}
\tag{26}
$$

4. Conclusion and discussion

Example of DC, IT capability indicators assesment and optimization problems examples discussed. The problem statement described can be used to formulate a number of mathematical problems statements for DC, IT capability research and related mathematical problems statements. Use cases of such indicators researc include choosing optimal IT, IC characteristics, digitalization planning, synthesis of information operation characteristics based on mathematical models of IT use for a system's functioning, and strategic planning based on the analytical investigation of DC use indicators. Further research will be directed on computational aspects of the discussed problem.

5. Acknowledgement

6. References

Aslaksen, E. W. 2009. Designing complex systems: Foundations of design in the functional domain / Erik W. Aslaksen, Boca Raton: CRC Press.

Davis, J. P., Eisenhardt, K. M., and Bingham, C. B. 2009. "Optimal Structure, Market Dynamism, and the Strategy of Simple Rules," Administrative Science Quarterly (54:3), pp. 413-452 (doi: 10.2189/asqu.2009.54.3.413).

Dechter, R. 2019. Reasoning with probabilistic and deterministic graphical models: Exact algorithms / Rina Dechter, San Rafael, California: Morgan & Claypool Publishers.

Simon, H. A. 1996. The sciences of the artificial, Cambridge, Mass: MIT Press.

THE ROLE OF THE FOURTH INDUSTRIAL REVOLUTION IN THE MODERN WORLD

Valentina Dzobelova

North-Ossetian State University named after Kosta Levanovich Khetagurov
dzobelova@mail.ru

Alisa Olisaeva

North-Ossetian State University named after Kosta Levanovich Khetagurov
alisa.olisaeva@mail.ru

Sergey Yablochnikov

Technical University of Communications and Informatics
vvkfek@mail.ru

Valentina Akasheva

National research Mordovian state university of N. P. Ogaryova
valakasheva@mail.ru

Sergey Gerasimov

Saint-Petersburg State University of Economics (UNECON)
votje82@mail.ru

Keywords

Industry 4.0., Industry 3.0., digital economy, the Fourth Industrial Revolution, modernization, information technology, smart factory, digital factory, virtual factory, stage.

Abstract

This article discusses the main stages of the transition of Industry 3.0 to Industry 4.0., which is also called the Fourth Industrial Revolution. The selected topic is undoubtedly quite relevant in connection with the fact that it touches on the problems of methods and methods for the development of information and communication technologies, automation and robotization of production processes in the modern world. The main technological units of the economies of the future are also noted, which are: smart, digital and virtual factories. The main types of digital technologies that are to be introduced in advanced industries are identified.

1. Introduction

Modern manufacturing enterprises implement their activities in accordance with the technological processes of Industry 3.0. New hardware and software and automation systems are being introduced into design and production activities that increase the productivity of specialists and reduce the time it takes to complete individual stages of the product life cycle.

An analysis of ways to modernize the enterprise's production capacity illustrates that the main tasks of enterprises are not only the reduction of production costs, automation, etc. but also a digitalization of design procedures, production processes, supply, etc. throughout the product life cycle (Abdrakhmanova, G. I., Gokhberg, L. M., Demyanov, A. V. and other, 2018, p.89).

The first concepts, as well as the concept of the modern digital economy, appeared at the end of the last century. In 1995, Nicholas Negroponte, an American scientist from the University of Massachusetts, introduced the concept of "digital economy". The advantages of the digital economy, in Nicolas Negroponte`s opinion, could be the following: the lack of physical weight of products, which is replaced by information volume, lower costs of resources for the production of electronic goods, a much smaller area occupied by products, and also the almost instantaneous movement of goods via the Internet.

Thanks to technological progress, the development of information and communication technologies and the Internet, digital technologies are penetrating our lives increasingly. Modern technologies are changing the world, so the authorities must keep up with the times to effectively regulate the economy: to develop programs adapted to current trends in the field of information technology.

Today, the electronic economy is already beyond the scope of commercial aspects (Babkin A.V., Burkaltseva D.D., Kosten D.G. . and Vorobyev Yu.N., 2017, p.10). The digital economy is being introduced in large companies and corporations, as well as in the social spheres of the life of the population of countries, besides, government organizations and structures are being introduced into the work (Annual analytical report, 2017).

In modern conditions, the growing importance of information technology and the digital economy is becoming increasingly important, thereby displacing the old norms in all areas of public policy. Over the past few years, we can already note significant changes in the lives of the population of our and other countries of the world along with changes in the distribution of jobs; many new professions have already appeared, associated with a technological breakthrough. (Dzobelova, V., Yablochnikov, S. and Semenova, L. 2019, p. 83).

In the practice of international comparisons the results of scientific and technological activities by indicators of patent activity, technological balance and the results of fundamental research are assessed using bibliometric indicators (Dzobelova, V.B., Olisaeva, A.V., 2018, p. 474).

The world is undergoing rapid development and implementation of the so-called Industry 4.0, which is based on the use of completely new digital technological processes from the design stage to the stage of product disposal. At the same time, Industry 4.0. includes: additive and cloud technologies; security technologies of cyber-physical systems; technology of "augmented reality"; Internet of Things technology (IoT); BigData technology for processing large volumes of production data; industrial sensoric etc.

Thus, Industry 4.0, which is also called the Fourth Industrial Revolution, is a set of technical and economic solutions that are implemented in the design and production activities of enterprises to organize an integrated environment for the development and production of products using digital information technologies.

2. Methods

The term "Industry 4.0." was first introduced in 2011, and is the name of one of ten projects of the state Hi-Tech Strategy of Germany, which reveals the essence of the concept of Smart Manufacturing using the base of the global industrial network of the Internet of Things and Services

It is the construction of production facilities to maintain an unmanned production and design environment for virtualization technology that are the main expectations from the use of Industry 4.0. technologies (Abdrakhmanova, G. I., Gokhberg, L. M., Caves, M. A. and other, 2017, p.20).

As a rule, any enterprise passes through six stages in case of transition from Industry 3.0. to Industry 4.0. in connection with the desire to improve its design and production infrastructure and introduce new technologies.

As a rule, computerization and networking (which are actively used in almost all enterprises of Industry 3.0.) are defined as the first two stages on the path of modernization of an industrial enterprise to Industry 4.0.

In modern conditions, the introduction of Industry 4.0 leads to an increase in the quality of life, in turn, leads to an increase in human capital, as it is one of its components (Olisaeva, A., Dzobelova, V., Yablochnikov, S., Cherkasova, O. and Davletbayeva, N., 2019, p.84)

In respect to the stage of computerization, at this stage, the design and production infrastructure of the enterprise is modernly equipped with the introduction of new computer-aided design systems for instrumentation.

The stage of networking involves the introduction of a wireless computerized control system based on IoT in industrial enterprises.

The other four subsequent stages, as a rule, are new for the enterprises of Industry 3.0., but these are the main stages of Industry 4.0.

That is, the "virtualization" stage, which consists in introducing technologies for using digital product models and "digital doubles" of production processes and the enterprise as a whole, is followed by the "networking" stage.

The fourth stage is called "transparency." It involves the introduction into the design and production environment of technologies for the collection, analysis and formation of management teams using methods and means of BigData processing of a significant amount.

The fifth stage is "forecasting". This means the use of various mathematical models, algorithms in the design and production environment. With their help, it is possible to carry out modelling of the processes of design, production, logistics and operation of products that exist only in the form of digital models.

And the final stage is the "adaptability" of the enterprise, accompanied by the introduction of artificial intelligence technologies focused on self-organization of flexible automatic production lines.

In the Russian Federation, the process of modernizing the production capacities of Industry 3.0 and the introduction of technologies of Industry 4.0. into the design and production environment is regulated by a normative act "On the Strategy for the Scientific and Technical Development of the Russian Federation", within the framework of which the road map of the National Technological Initiative "TechNet" functions. The main purpose of the "TechNet" card is to improve gradually the production capacities of enterprises and create certain conditions for the digitalization of domestic enterprises.

Modernization of production capacities of Industry 3.0. in our country will lead to the formation of three new types of enterprises of Industry 4.0., which are commonly called the "factory of the future", which includes: smart, digital and virtual enterprises.

By type of design and production activity this distribution of "factories of the future" is associated with technological processes taking place at these enterprises.

3. Analysis

The main stages of a product's life cycle related to design are landmarks of a "digital enterprise". Using the analysis of the roadmap "TechNet" one can note that it is an analogue of the "digital enterprise" of Industry 4.0. in Industry 3.0. is an enterprise with a format of research institutes and experimental design bureau.

 "Smart" enterprises are mainly focused on the technological processes of manufacturing the product. As a basis for their activities, the developers of the "TechNet" roadmap put forward the idea of creating unmanned production, the functioning of which will be based on digital models and digital doubles of the product and the enterprise as a whole.

The main goal of the "virtual enterprises" is to merge the design and production processes of both the "digital enterprise" and the "smart" to support the adopted management decisions in the product life cycle.

The table below shows the aggregate market of Factories of the Future (FF) in the world, as well as an assessment of the technology markets on which the services of the participating companies of "TechNet" will be based. The priority market in this context is the engineering and design services market.

Table 1 - Characteristics of Markets of Factories of the Future

Markets of Factories of the Future (thousand US dollars)	2015	2020	2035
Engineering and design services market	World – 773 000 00011 RF – 2 200 00012 Share of the RF - 0,28%	World – 896 118 859 RF – 2 509 133 Share of the RF– 0,3%	World – 139600000013 RF – 10900000 - 21200000 Share of the RF-0,8-1,5%
Accelerated certification system and services market	World – 0 RF - 0	World – 5 000 000 RF – 25 000 (0,5%)	World – 33 600 000 RF – 160 000 (0,5%)
Educational services market	RF – 8 000	RF – 13 861	RF – 50 000- 94 000
Global technology markets for the FF components	368 300 000	478 400 000	1 757 600 000
incl. Digital Design and Modeling (CAx, PLM, PDM, etc.)	23 900 000	31 500 000	74 800 000
incl. Numerical control machine tools	54 500 000	94 200 000	281 400 000
incl. Additive technologies	5 200 000	18 100 000	216 400 000
incl. Hardware support (including HPC)	4 800 000	7 200 000	24 300 000
incl. Advanced materials	63 700 000	88 800 000	145 400 000

incl. Robotic devices	22 000 000	48 900 000	241 600 000
incl. Manufacturing execution system (MES) and Industrial control system (ICS)	127 000 000	180 300 000	366 100 000
incl. Enterprise information management systems	25 600 000	39 300 000	92 600 000
including Big Data	29 200 000	38 700 000	90 000 000
including Industrial Internet	12 400 000	25 600 000	225 000 000

Support for the development of advanced manufacturing technologies by states and their associations, as well as by industrial companies - world leaders, is as follows: a) the Advanced Manufacturing Partnership initiative in the United States since 2011 and the creation of institutes of manufacturing innovation in the field of additive technologies (America Makes, 2012) – Digital Manufacturing and Design Innovation Institute (DMDII, 2014), Institute for Advanced Composites Manufacturing Innovation (IACMI, 2015), Lightweight & Modern Metals Manufacturing Innovation Institute (LM3I, 2014). b) Public-Private Partnership (PPP) for the creation of Factories of the Future in the European Union (2008, restarted as part of the Horizon 2020 programme in 2014) and a strategy for the development of key promising technologies (2009) c) the Industrie 4.0 Project in the framework of the Germany`s High-Tech Strategy 2020 Action Plan (2012). d) the Made in China 2025 Plan (2015). e) models of Digital Enterprise and Digital Factory by Siemens, Brilliant Factory by General Electric, etc.

For the Russian economy, 2019 was a turning point in a sense: large potential consumers of new technological developments intensified, large investors (state corporations) joined to the modernization of production. The beginning of the strategic shift was the approval of the national project Digital Economy, the cost of which by 2024 should be more than 1.6 trillion rubles (The federal target program "Digital Economy of the Russian Federation Program", 2017). Last year, the development of new technologies for the first time received the status of a full national priority - digitalization was actively discussed at all levels, with officials and large businesses competing in the degree of enthusiasm. The passport of the national project "Digital Economy" was approved in December 2018, last year the previous government formed the main architecture of support measures, and state-owned companies entered into agreements on the development of several priority technological areas. In 2019-2021 about 403 billion rubles are allocated for the implementation of the national program "Digital Economy". Of these, about 108 billion rubles will be spent in 2019, a little more than 123 billion rubles - in 2020, and almost 172 billion rubles - in 2021. In addition to federal investments, it is planned to attract financing of projects in the field of digitalization from business and regional authorities.
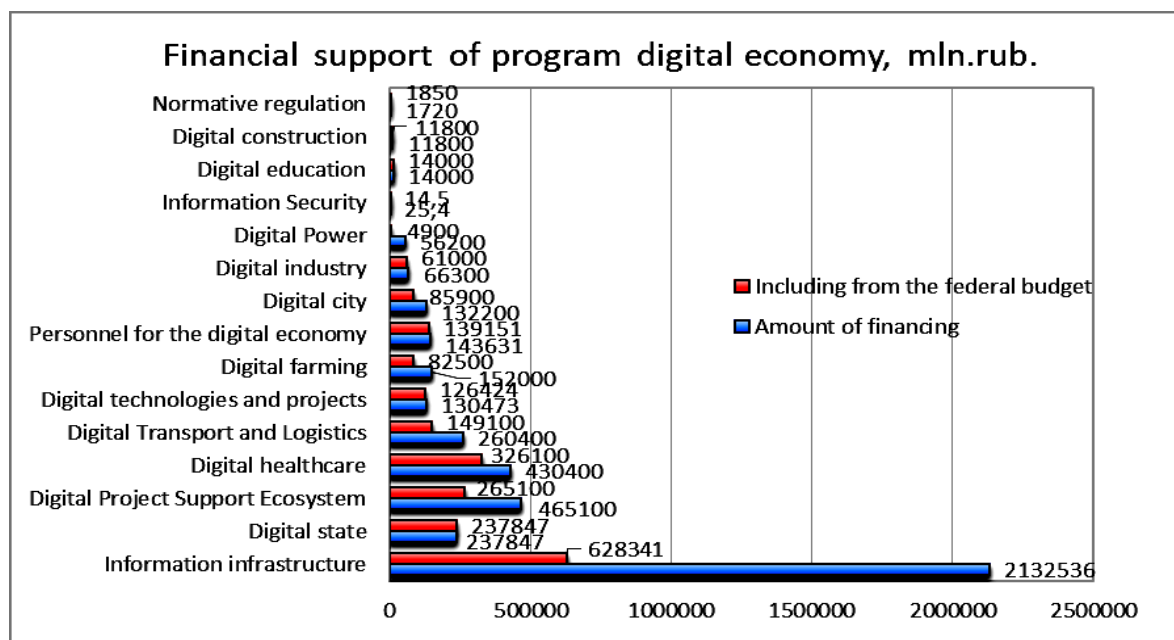
Figure 1 - Cost chart for the implementation of the national program "Digital Economy" in 2018-2024

Figure 1 shows the cost chart for the national program "Digital Economy of the Russian Federation", following the project submitted by the Ministry of Digital Development to the government in August 2018.

In general, federal budget spending on the national project "Digital Economy" by 2024 will amount to 1 trillion 80 billion rubles. These costs are necessary to solve important tasks, one of which is to improve the investment climate in the country, the program will have a significant impact on the development of Russian small and medium-sized businesses. Today, dramatic changes are taking place in the field of information and communication technologies, which are reflected in various areas of company activity. The Internet enables even new and tiny companies to sell their products around the world. Companies can emerge and grow quickly, with relatively little capital investment.

 The market position of companies in the digital economy is becoming increasingly complex; risks and the level of uncertainty in making strategic decisions are growing. This situation is associated with unstable conditions due to dynamic changes at the technological level, increased competition, and the influence of the state on the economy.

Last December, state corporations also presented their plans for the development of digital technologies. In July 2019, Sberbank became responsible for artificial intelligence, "Rostec" - for quantum sensors, blockchain and the Internet of things, "Rostelecom" and "Rostec" - for wireless communication technologies (5G), "Rosatom" - for quantum computing, and Russian Railways - for quantum communications. According to the results of stress tests of seven "cards", two of them (according to quantum sensors and blockchain) were sent for "deep study". In total, seven "roadmaps" have been approved: distributed registry systems, wireless communication technologies, components of robotics and sensors, virtual and augmented reality technologies, neurotechnologies and artificial intelligence, quantum technologies and new production technologies.

State-owned companies will not receive funds directly from the budget but will be able to influence the allocation of support by development institutions.

There will be three forms of support: participation in the capital, subsidization of interest rates and grants. The network of support institutions will include the Innovation Support Fund, RVC, VEB Ventures, Rusnano, the Russian Fund for the Development of Information Technologies, RDIF and

"Skolkovo". Also, one support measure (direct subsidization) will be assigned to the Ministry of Digital Development, Communications and Mass Media of the Russian Federation. Based on the above stages for the transition from Industry 3.0. to Industry 4.0., it is necessary to review the role of components of the main and auxiliary infrastructure of production, in particular, the role of technological equipment as a cyber-physical system, which is carried out by a computerized production management system.

There should also be mentioned the existing restrictions in the implementation of the "TechNet" such as:

1. Legislative limitations are caused by the peculiarities of the legislative systems of different countries, which are pioneers in the development and integration of technological solutions in the field of advanced manufacturing technologies (AMT). The experience of such countries as the USA, Germany, Great Britain, the Netherlands, Japan, South Korea, the People's Republic of China (PRC) and others should be considered when making an effective legal and regulatory framework for the development of AMT in Russia. The main legislative limitations faced by leading countries in the field of AMT include:

- laws and standards governing the process of the relationship between a person and a new generation of cyber physical systems (including human- computer interaction interfaces);

- regulation of the activities of key industry entities acting through the aggregator companies that act on the market as many small entities, representing them as a whole; and their activities at the moment remain outside the regulations in many aspects;

- the legal and regulatory framework in the field of development, prototyping and integration of package solutions, including technologies, standards, programs and systems between the solution package provider and the consumer.

2. Financial constraints are caused by the following factors:

- the difficult economic situation in the country, affecting the investment opportunities of both private and a public sector;

- under the conditions of economic uncertainty, financial entities are not tended to provide financing for high-risk assets, which often include innovative projects;

- an insufficient number of successful projects in the field of the pilot implementation of new or complex technological solutions based on the existing industries determines a general conservative point of view on the part of a business and the state regarding the financing of new projects in this area.

3. Infrastructure restrictions in the country are associated with the existing breaks in the innovation cycle (basic and applied research – experimental development – testing and implementation of complex technical solutions), as well as with the absence or deficiency of a number of world-class competencies for the implementation of the technological package necessary for the implementation of the "TechNet" arrangements. Despite the presence of many elements of the innovation infrastructure, it is necessary to create mechanisms for their coordination to ensure the transition of innovations from stage to stage, thematic focus and the results` conjugation, the practical application of the developed solutions in Russian and foreign markets.

4. Conclusion

Also, the role of specialists who perform production functions in enterprises with Industry 3.0. is subjected to the review, since with the introduction of technologies of Industry 4.0. a person becomes

secondary since a cyber-physical system oriented to service begins to play his role. All this leads to the transition of workers from the main production system of Industry 3.0. to the auxiliary production system of Industry 4.0.

The demand for the TechNet integrated technological solutions will be formed, first of all, in promising markets in growing sectors of the economy due to the formation of additional resources for their implementation, as well as the need to maintain the rate of the increase of production capacities, including the creation of new ones. According to the Center for Macroeconomic Analysis and Short-Term Forecasting, (CMASTF), several economic sectors will demonstrate significant growth in output by 2035.

The degree of achievement of the goals of scientific and technological development depends on the implementation of conditions such as: orientation to global competitiveness at all stages of creating knowledge and value, which involves taking into account global trends, increasing the efficiency and effectiveness of the field of science and technology, including through the transition to its new models, strengthening human resources and material and technical base; the presence of a flexible, receptive innovation ecosystem for the technological modernization of existing and development of new industries; stimulating demand for scientific and technological achievements; improving the quality of scientific, technological and innovative policies, creating a system for assessing the effectiveness of measures implemented, ensuring their adaptability to global trends, including the effective use of alternative institutions that meet the specifics of the digital environment.

Thus, we can note the important role of Industry 4.0., which is of great importance for the whole world. The task of any country is to increase efforts to raise the country to a new level because great attention is being paid to the digital economy in the modern world.

5. References

Abdrakhmanova, G. I., Gokhberg, L. M., Caves, M. A. and other. (2017). Indicators of digital economy: 2017. Textbook for Econ. spec. universities, Higher School of Economics, 1-320.

Abdrakhmanova, G. I., Gokhberg, L. M., Demyanov, A. V. and other. (2018). Digital economy: a brief statistical compilation. Textbook for Econ. spec. universities, Higher School of Economics, 1-96

Annual analytical report "Runet. Runet Economy. Digital economy". Report of the Association of electronic communications (RAEC), 2017.

Babkin A.V., Burkaltseva D.D., Kosten D.G., Vorobyev Yu.N. (2017). Formation of the digital economy in Russia: the essence, features, technical normalization, development problems. Scientific and Technical Reports SPbSPU. Economics, 10(3), 9-25.

Dzobelova, V.B., Olisaeva, A.V. (2018). Analysis of innovative development of the NCFD regions in Russia. IDIMT 2018: Strategic Modeling in Management, Economy and Society - 26th Interdisciplinary Information Management Talks, 473-479.

Dzobelova, V., Yablochnikov, S., Semenova, L. (2019). Practical application analysis of information and communication technologies in the socio-economic sphere. ACM International Conference Proceeding Series, 80-85.

Olisaeva, A., Dzobelova, V., Yablochnikov, S., Cherkasova, O., Davletbayeva, N. (2019). Formation and development of the digital economy in modern conditions - Development within the framework of industry 4.0. IDIMT 2019: Innovation and Transformation in a Digital World - 27th Interdisciplinary Information Management Talks, 83-88.

The federal target program "Digital Economy of the Russian Federation Program" was adopted by the Resolution of the Government of the Russian Federation on July 2017.

INNOVATION, NEW BUSINESS MODELS AND STRATEGIES

INNOVATIONS FOR FUTURE ENERGY SECURITY

Tomáš Pitner

Masaryk University, Czech Republic
94@muni.cz

Jan Ministr

VSB Technical University of Ostrava, Czech Republic
jan.ministr@vsb.cz

Keywords

Abstract

Energy sector faces a unprecedent fast growth while undergoing substantial change due to the necessity of climate change mitigation measures such as increasing the role of renewables. The paper presents a key role of IT in substantial innovation of energy sector, from production, transmission, distribution, consumption, and the business around. The trend towards decentralization and creation of new and local markets would not be possible without novel technologies such as blockchain that are robust and introduce decentralized trust. Together with ageing or missing infrastructure in developed or developing countries, respectively, the network capacity reserve is getting a bottleneck which can be eliminated by network-safe usage of flexibility.

1. Introduction

1.1. Energy Sector Today and Tomorrow

Energy sector, namely electrical energy, is a critical component of current economics and society. Its significance is even expected to grow in the coming decades. According to New Policies Scenario of International Energy Agency (IEA, 2018): *"Under current and planned policies, modeled in the **New Policies Scenario**, energy demand is set to grow by more than 25% to 2040, requiring more than $2 trillion a year of investment in new energy supply."*

From the point of view of a country where power grid has connected virtually all even small places across the country almost a century ago, it is not imaginable that there are still regions on Earth without electricity supply. Thus, this aspect must also be taken in account for future considerations.

Electricity shows it potential to accelerate the transition towards sustainable energy supply being covered by increasing proportion of renewable energy sources. Not only the use of renewables but also other technological advances, such as emerging energy storage technologies, flexible generation and co-generation of electricity and heat promise to eliminate negative impact of extensive energy usage going hand-in-hand with growing population and increasing living standards bringing new opportunities and expectations of the world population.

Though many of such activities and technology changes are initiated bottom-up by businesses and individuals, governments play a key role in power sector being both critical infrastructure and strong network-type industry requiring strategy planning, decision making, coordination, and even subsidies in order to progress and guarantee secure, dependable and sustainable future.

As IEA indicated in its Annual Report (2018): *"Major transformations are underway for the global energy sector, from growing electrification to the expansion of renewables, upheavals in oil production and globalization of natural gas markets. Across all regions and fuels, policy choices made by governments will determine the shape of the energy system of the future."* So, the role of governments, infrastructures, and system-wide is crucial.

1.2. Goals of the Paper

The paper is divided into two main parts – primarily, it gives a structured overview of the main challenges we as a modern society must cope with. It concentrates at key technological and business innovations bringing new opportunities for reaching sustainable power supply to cover growing needs. Among them, security is a key prerequisite which is true for both physical- and cyber-security of power production and grids. So many innovations focus at ensuring reliability and security of electricity supply and trading.

The latter part (Chapter 4) focuses at our research in power grid flexibility identification, modelling, and exploitation which is an innovative way reacting on increasing variability of electricity supply from volatile renewable sources such as wind and solar.

In general, the paper serves as an overview and outlook into the energy sector as a promising playground for many future innovations that are by far not limited by core technology but range from basic technology for production, storage, balancing, conversion, and use to business models e.g. for online fast trading.

1.3. Methodology

The methodology behind this paper is based on a literature meta-review (Chapter 2, 3) accompanied backed by own research contributions (Chapter 4) leading to construction of several new models of flexibility being capable to be established on today's Czech electricity market while being compatible with emerging pan-European energy market regulations and platforms.

2. Challenges and Opportunities in Energy Sector

The IEA has identified following top threats and opportunities forming the energy sector now and in coming decades. Shortly, the main issues arise from:

- *Increasing share of renewables* – there are already large countries leading the peloton: Germany, California, Norway, Denmark where renewables represent 20-40 % of energy portfolio while planning to reach 50 % and more soon, targeting zero-carbon economies till the middle of century (Martinot, 2016).

- *Increasing Affordability of Renewables* – current rise of renewables would never be possible at current pace until the technology itself became affordable: the prices have already dropped significantly and will continue to drop by 90 % (solar) and 70 % (wind) till 2040 as forecasted by IEA (2018).

- *Growing Significance of Electricity vs ageing Infrastructure* – not only the total demand for electricity increases but also its share within the energy mix. This trend will continue as

completely new sectors get converted to electricity, such as mobility (cars, bikes, but also trucks, trains, and even airplanes), increased demand for air conditioning due to rising global temperatures as well as further growth of production in manufacturing. This trend, however, faces challenges of insufficient and/or retiring infrastructure where the lifespan of decades (30-40 years for major lines) does not correspond to current fast growth.

- *Developing economies* – they will represent the largest portion of the growth in demand. Due to their underinvestment so far, if they continue to rise, they create the major increase while having substantial problems with their electricity grids, their stability, wide availability of electricity, and suboptimal share of renewables as they still largely depend on fossil fuels.

3. Energy- and Cyber-security Challenges

3.1. Innovative Energy Storage Facilities

Temporarily superfluous electric energy can nowadays be stored in facilities based on *mechanical, electrochemical, chemical, thermal,* or *electromagnetic* energy storage principles. In contrast to energy sources, storage facilities allow to load and supply energy always in a limited quantity given by the technology and size of the facility.

According to Acar (2018), main criteria of efficient energy storage encompass:

- *High storage capacity* effectively means high energy density (measured in kWh/kg) in order to reach reasonable volumes of energy to be stored while keeping a feasible facility size;

- *Low cost* (both CAPEX and OPEX) means the technology should not be excessively demanding in both development and operation;

- *Adaptability* it must support growing demand for flexibility;

- *High efficiency* must be ensured in order not to negate the positive effects;

- *Better environmental performance* necessarily outperforming conventional power stations.

Different technologies imply their different positions at the scale between high capital costs (and low operating costs) and cheap equipment with high operating costs. While the former is suitable for long-term energy storage, such as season-to-season to balance production from renewables, eg. photovoltaic energy produced during sunny summer days to be consumed dominantly in winter, or to balance between windy and quiet days when using wind turbines, the latter sources are suitable for urgently needed *restoration reserves* or *balancing* where one can afford more expensive but quick energy although available in smaller quantities.

To quickly compare performance, Acar (2018) constructed a simple model enabling comparison of substantially very different technological principles. Taking all dimensions into account, Acar's performance model can be used to assess an overall performance of energy storage technology:

- For mechanical energy storage systems, it is *pumped hydro*;

- For electrochemical energy storage systems, it is *high temperature batteries*;

- For chemical energy storage systems, it is *ammonia*;

- For thermal energy storage systems, it is *molten salt*;

- For electromagnetic energy storage systems, it is *superconducting magnets*;

In general, average performance rankings of mechanical, electrochemical, chemical, thermal, and electromagnetic energy storage systems are *growing from electrochemical (lowest) to chemical (highest)*.

3.2. Blockchain in Energy Trading

Blockchain (or originally *block chain*) was first published by an author Satoshi Nakamoto (whose real identity has not been revealed since then) in 2008. It serves as a cryptographically based public transaction ledger, originally for cryptocurrencies but quickly expanding to other application domains. The main advantages are independence of a central authority (server) keeping the records, and – in contrast to older cyber-currency models – principal avoidance of double-spending.

Energy trading is a major application of blockchain along with financial markets, video games, and supply chain management. According to British government study, blockchain shows the potential to '*reform our financial markets, supply chains, consumer and business-to-business services, and publicly-held registers*', see Walport (2016).

Mangelcamp et al (2017) and Andoni (2019) brought up a thorough overview about the state-of-the-art of blockchain applications in either local energy trade in small grids as well as on large markets. The main areas identified as opportunities include: *emerging peer-to-peer (P2P) energy trading and Internet of Things (IoT) applications, to decentralised marketplaces, electric vehicle charging and e-mobility,* see Andoni et al (2019). It is obvious due to the technology features that blockchain position will be strong wherever we can benefit from the distributed and robust nature without a single point of failure, see (Nyemkova et al, 2019).

Since the future energy systems can be characterized by three key principles: *decarbonisation*, *decentralisation* and *digitalisation*, digital technology is a major pillar with blockchain helping to establish completely new markets enabling energy trade with even small players such as „prosumers", i.e. small clients participating as both consumers and producers (e.g. having photovoltaics panels, or being able to provide their electric car capacity for charging/discharging on demand).

Another advantage of bottom-up peer-to-peer energy market sis their potential to reduce unnecessary costs generated from the decoupling between whole-sale and retail prices on traditional markets which is estimated to be 1.4+ billion GBP solely in the UK, see (Business Energy prices, 2020).

The blockchain can change the energy business according to Andoni et al (2019) the following ways:

- *Billing*: smart contracts and smart metering can realise automated *billing* for prosumers. Utility companies might benefit from the potential for energy *micro-payments*, pay-as-you-go or pre-paid business models;

- *Custom-tailored tariffs*: blockchain together with AI can help to create develop *tailor-made custom offers* dependent of invididual preferences and energy usage profiles.

- Blockchain can help *green certificates* trading.

- *Local energy markets* can be established which would help to *reduce technical losses* induced when transmitting the locally produced energy across higher voltage levels to distant consumers instead of being sold and consumed locally.

- *Local markets* can be supported by *shared resources* such as community storage being managed with help of blockchain.

- *Smart metering and smart grid* can benefit from the *trust* introduced by blockchain in building and grid monitoring.

- *Confidentiality, Identity management & Privacy* are serious issues in current smart grids where blockchain can help significantly (Nyemkova et al, 2020).

Moreover, blockchain promises substantial simplification of whole-sale energy market relations when properly applied. The following Figure depicts the future architecture.
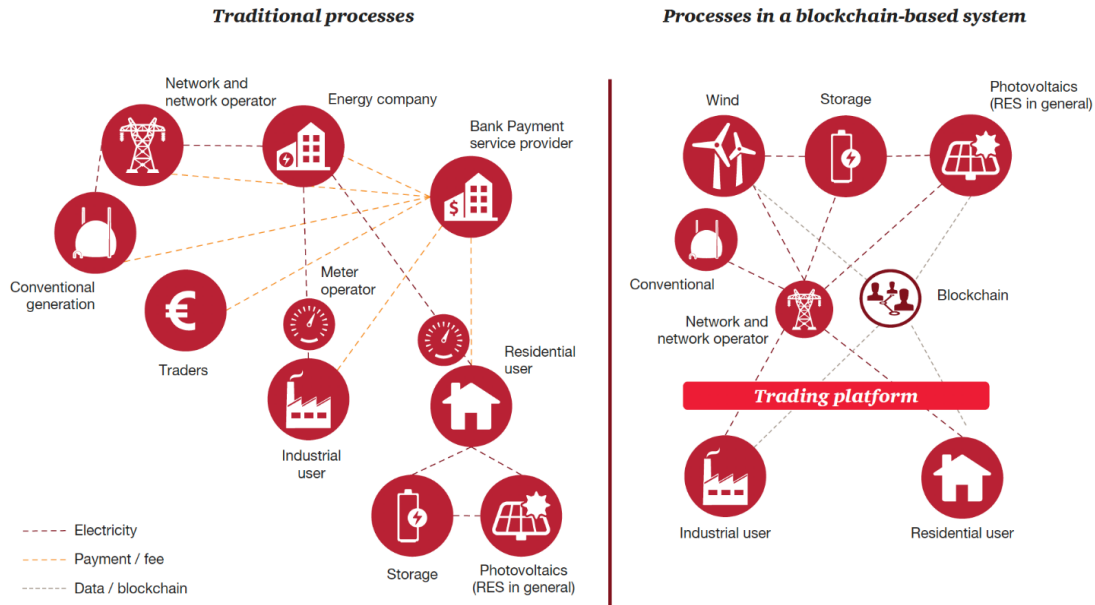
Fig. 1 Energy market architecture by blockchain (PWC, 2016)

Intra- and cross-border imbalance settlement is nowadays a lengthy process. According to Elexon (2017), settlement takes up to 28 months to finalize in UK. Employing blockchain could reduce the time to minimum while saving credit needed and reduce risks. At the same time, quick transactions allow to react upon the market, e.g. increase or reduce demand according to price signals.

Distributed decision-making on energy markets can be enabled by blockchain and IoT. Since blockchain is ready for M2M transactions, the number of smart devices rises rapidly, and availability of fast communication grows with the advent of 5G networks, we can expect applications of smart metering far beyond meter reading and billing. It can be used for network assets health diagnostics, measuring quality of energy delivery (SAIDI and SAIFI indicators) and the like. It shall enable complex integration of intelligent metering and regulation systems in residential or office building, factories, as well as local grids and microgrids.

Blockchain can become an instrument changing digital economy as we know it (Doucek & Hološka, 2019). In energy sector, its impact is expected to be specifically strong.

4. Flexibility

4.1. Definition

According to Denholm and Hand (2011) *"System flexibility can be described as the general characteristics of the ability of the aggregated set of generators to respond to the variation and uncertainty in net load. At extremely high penetration of variable generation, a key element of system flexibility is the ability of baseload generators, as well as generators providing operating reserves, to reduce output to very low levels while maintaining system reliability."* or, alternatively, *"the ability of devices (consuming, producing or storing electricity) to change in response to a command or price signals the amount of (consumed or supplied) energy compared to the agreed/assumed diagrams"*.

Due to evolution of electricity demand and supply towards more dynamic, uncontrollable, and to certain extend unpredictable, e.g. photovoltaic or wind sources, greater flexibility is needed to manage peaks and gaps in demand or supply.

4.2. Flexibility Needs for Grids

There is an estimation to meet 25% renewable energy target with minimal grid system changes for selected ASEAN nations, see (Huang, Kittner & Kammen, 2019). We can observe similar figures for the U.S. where variable renewable energy at 30% is possible with minimal system changes to regional grids (Cochran et al, 2015).

4.3. Production Control to Reduce Flexibility Needs

- *Geographical distancing* - some flexibility requirements can be naturally reduced by reducing variability by exploitation of anticorrelation of certain renewable sources. Typically, more geographical diversity is needed, e.g. photovoltaic in more distant regions are less likely to have the same drop in production due to local clouds.

- *Anticorrelation* – wise combination of different resources such as solar and wind may reduce fluctuations by mutual compensation – during cloudy days with less sunshine, it is more likely that there is stronger wind helping to compensate it.

- *Orientation of photovoltaic* – if the solar facility allows to change East-South-West orientation of solar panels (or W-S-E on Southern hemisphere, analogically), it makes the solar production more balanced throughout the day.

4.4. Energy Storage to Reduce Flexibility Needs

There are technologies that can contribute to grid flexibility by storing energy. They can be classified by physical principles and impact thereof on storage periods (Ajanović, 2020):

- *Constant renewables* (e.g. geothermal) – allow (partially) flexible exploitation thus regulating energy production;

- Renewables allowing *fuel storage* (e.g. biomass, biogas) – partially allowing storage and changes in supply even across seasons;

- *Utility-scale energy storage* (e.g. pumped hydro, molten salt) – enabling short-term storage to balance supply-demand during short periods: hours and days;

- *Fuel storage* (e.g. hydrogen gained by hydrolysis and consumed by power cells) – allowing mid-term to long-term storage of generated fuel;

- Small-scale direct *electricity storage* (e.g. batteries, compressed air, flyweels) – usually for short-term storage serving for special purposes, e.g. auxiliary services for transmission systems;

- *Demand-response* (e.g. smart appliances allowing consumption peek shaving using time shifting, plus precharge and/or rebound) – is an opposite approach; instead of shifting supply, it shifts demand or shaves peaks in demand.

4.5. Use of Network-Safe Flexibility

The project SecureFlex is part of long-term academic-industrial collaboration (Ministr & Pitner, 2016). It aims to create analytical, computational and optimization tools and targeted studies that will

contribute to a systemic energy solution for the safe use of power flexibility associated with the integration of new technologies and market players in the Czech Republic. The timing of the project corresponds to the expected dates of implementation of the so-called *Winter package* of the EC.

We distinguish two basic types of flexibility (Mamula et al, 2019):

- *"Energy Trading" flexibility* – where the agreed volume of energy is delivered in the agreed trading interval. The resulting product is traded on standard whole-sale power markets. The volume of energy is traded without further requirements for the shape of the diagram and the payment for this type of flexibility is only a payment for energy (volume). Provider of Flexibility (POFL) may require "preload" or "rebound" after flexibility is provided (typically systems with storage capacity or hysteresis characteristics).

- *"Power Trading" or "auxiliary services" flexibility* – reserve range for activation is made available for timeframe where activation may occur at any time within the agreed time interval of availability, but it does not have to occur. Characteristics of the diagram include also the ramp-up and ramp-down edges (gradient). Payment for this type of flexibility includes payment for supplied control energy and payment for availability (reservation).

The project SecureFlex has developed a flexibility provisioning and trading concepts introducing applicable aggregation models with new market roles, namely providers of flexibility and flexibility aggregator.

5. Conclusion

Energy supply, demand, as well as grids recently undergo a substantial change which will continue even at a faster pace. We will probably see continuing growth of innovations in energy production namely from renewables, long-term effective storage, but also increasing use of flexibility, implementation of local- and microgrids, and completely new (distributed and small) market models. Some of these concepts could never be implemented without strong IT support based on blockchain and similar "distributed trust" technology, namely the ability of fast transaction processing without a single point of failure. Flexibility in power grids can only be achieved using smart metering and smart grid mechanisms providing enough data and decision-making capabilities.

6. Acknowledgement

7. References

Acar, C. A comprehensive evaluation of energy storage options for better sustainability, International journal of energy research Volume: 42 Issue 12 (2018) ISSN: 0363-907X Online ISSN: 1099-114X

Ajanovic, A., Hiesl, A., & Haas, R. (2020). On the role of storage for electricity in smart energy systems. Energy, 117473. https://doi.org/10.1016/j.enpol.2019.01.025

Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., ... & Peacock, A. (2019). Blockchain technology in the energy sector: A systematic review of challenges and opportunities. Renewable and Sustainable Energy Reviews, 100, 143-174.

Business Electricity Prices. (2020). Available at https://www.businesselectricityprices.org.uk/retail-versus-wholesale-prices/

Cochran, J., Denholm, P., Speer, B., & Miller, M. (2015). Grid integration and the carrying capacity of the US grid to incorporate variable renewable energy (No. NREL/TP-6A20-62607). National Renewable Energy Lab.(NREL), Golden, CO (United States).

Denholm, P., & Hand, M. (2011). Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. Energy Policy, 39(3), 1817-1830.

Doucek, P.; Hološka, J. Digital economy and industry 4.0 (2019) IDIMT-2019. Innovation and Transformation in a Digital World. 27th Interdisciplinary Information Management Talks (Sept. 4-6, 2019, Kutná Hora, Czech Republic). TRAUNER Druck GmbH & Co KG, Linz, 2019, 33-39.

Elexon. (2017). The electricity trading arrangements – A beginner's guide. Available at https://www.elexon.co.uk/wp-content/uploads/2015/10/beginners_guide_to_trading_arrangements_v5.0.pdf

Huang, Y. W., Kittner, N., & Kammen, D. M. (2019). ASEAN grid flexibility: Preparedness for grid integration of renewable energy. Energy policy, 128, 711-726.

International Energy Agency (2019). World Energy Model - Scenario analysis of future energy trends. Available at https://www.iea.org/reports/world-energy-model

International Energy Agency (2018). World Energy Outlook. Available at https://www.iea.org/reports/world-energy-outlook-2018

Mamula, O., D. Hrycej, M. Střelec, T. Pitner, F. Kasl, V. Stupka, F. Procházka. Flexibilita známá a neznámá. All for Power, Praha: AF POWER agency, a.s., 2019, roč. 13, č. 4, s. 57-61. ISSN 1802-8535.

Martinot, E. Grid Integration of Renewable Energy: Flexibility, Innovation, and Experience. Annual Review of Environment and Resources 2016. 41:223–51

Ministr, J., & Pitner, T. (2016). The Academic-Industrial Collaboration as an Innovation Instrument. IDIMT-2016: Information Technology, Society and Economy Strategic Cross-Influences: 24th Interdisciplinary Information Management Talks, Sept 7- 9, 2016, Poděbrady, Czech Republic.). TRAUNER Druck GmbH & Co KG, Linz, 2016.

Nyemkova, E., Chaplyha, V., Ministr, J., Chaplyha, V. (2020). Advanced Methods and Means of Authentication of Devices. Lecture Notes on Data Engineering and Communications Technologies: Business Information Data-Centric Business and Applications. Cham: Springer, pp. 205-228, ISBN 978-3-030-35648-4.

Nyemkova, E., Vojtusik, S., Ragulin, A., Chaplyha, V. (2019). Real-time authentication of electronic devices in cyber-physical systems. In Proceedings of 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T). Cham: Springer, pp. 589 – 592.Mengelkamp, E., Notheisen, B., Beer, C., Dauer, D., & Weinhardt, C. (2018). A blockchain-based smart grid: towards sustainable local energy markets. Computer Science-Research and Development, 33(1-2), 207-214.

Pitner, T., V. Stupka, F. Kasl, O. Mamula, D. Hrycej, M. Macaš. Analýza uplatňovaných motivačních schémat a modelů a jejich efektivity TK01030078-V18. Research Report of SecureFlex, 2019. 48 pp.

PwC global power & utilities. (2016). Blockchain – an opportunity for energy producers and consumers? Available at www.pwc.com/utilities

The Economist. (2015). The great chain of being sure about things. The Economist. 31 October 2015. https://www.economist.com/briefing/2015/10/31/the-great-chain-of-being-sure-about-things

Vyas, K. (2019). 7 Energy Efficiency Innovations Changing the Game. Interesting Engineering. 2 April 2019. https://interestingengineering.com/7-energy-efficiency-innovations-changing-the-game

Walport, M. G. C. S. A. (2016). Distributed ledger technology: Beyond blockchain. UK Government Office for Science, 1, 1-88.

INNOVATION OF THE ENDPOINT SECURITY SYSTEM

Jan Ministr

VSB Technical University of Ostrava, Czech Republic
jan.ministr@vsb.cz

Tomáš Pitner

Masaryk University, Czech Republic
tomp@fi.muni.cz

Vyacheslav Chaplyha

Lviv National Agrain University, Ukraine
4vyach@gmail.com

Keywords

Cyber security, cyberthreats, cybercriminals, cryptomining, cryptojacking, endpoint, Endpoint Detection and Response (EDR), ransomware, zero-day attacks

Abstract

Information system security is a crucial issue for a company's management today. Endpoint security is the practice of securing endpoints or entry points of end-user devices such as desktops, laptops, and mobile devices from being exploited by malicious actors and campaigns. The ensuring of cyber security is now increasingly costly for organizations. Malware threats are becoming more complex. 87% of IT managers agree that organizations lose an average of seven working days per month identifying and repairing infected computers (Sophos, 2018). In organizations, there is growing pressure to provide adequate endpoint security solutions. The article deals with problematic of the process of making endpoint security innovation decisions.

1. Introduction

The innovation of cyberthreats function as a system, so it is necessary to first understand the cause of the security threats that we want to eliminate (Franke, 2016). Cybercriminals use several techniques in their coordinated attacks, which usually include the following steps:

- *initial attack* - (phishing email that includes a malicious URL, clicking on a malicious link does not connect you to a command and control center - it downloads an initial malware, which then connects to a command and control server;

- *second attack* that using a combination of credential theft, privilege escalation, and malicious executables;

- *ultimate goal*, which could be stealing your data, or holding your data for ransom.

Only 16% of information security executives today are able to collect, analyze, and respond to 75% or more of their telemetry security events today (Forbes, 2019). The way to fight these attacks are coordinated complex implementation an interconnected cybersecurity system with integrated products that work together to overcome today's hackers. An effective integrated information system contains four basic elements:

- *central management* that allows to see and control everything in one place;

- *integrated components* they represent any different elements working harmoniously together;

- *automated actions* they provide sequential behavior based on pre-agreed criteria;

- *extensibility* that the security system can grow with increasing requirements.

These four features are what transform products into a system and are the basis for successful innovation (Danel et al, 2015). The stronger each of these features, the stronger the system (Danel & Neustupa, 2016).

2. Endpoint security innovation

Endpoint security is approach to secure entry points of end-user devices such as desktops, laptops, and mobile devices from being exploited by malicious actors and campaigns (Rozehnal &Danel, 2019). Endpoint security systems protect these endpoints on a network or in the cloud from cybersecurity threats.

Endpoint security innovation can include a variety of approaches. It is important for the end user to understand the threats they are trying to prevent. The procedure for selecting an endpoint security system can be divided into the following steps.

2.1. Assessment endpoint threats

The endpoint threat area is constantly evolving. It is important to consider the key endpoint threats when evaluating an endpoints security innovation, which are (Sophos, 2019):

- *Portable executables,* where malware protection is very important because it includes known and never-before-seen malware. The solution should be proficient in finding dirty and polymorphic files that have been modified to make them more difficult to identify Maksymovych et al., 2019).

- *Potentially unwanted applications (PUA).* Detection of PUA has become increasingly important with the rise of cryptomining programs used in cryptojacking attacks.

- *Ransomware,* which can be divided in two types: file encryptors and disk encryptors.

- *Exploit,* which is-based on file-less attacks are based on leverage techniques to take advantage of software bugs and vulnerabilities in order gain access and control of end user computer.

- *Active adversary techniques,* where many endpoint attacks involve multiple stages and multiple techniques. Examples of active adversary techniques include privilege escalation (methods used by attackers to gain additional access in a system), credential theft (stealing user names and passwords), and hiding malicious code inside legitimate applications.

2.2. Assessment of option the techniques of endpoint protection

Antivirus solutions are proven to be very effective against known threats. But there are two groups of techniques which we can use in innovation:

- *Traditional techniques* with endpoint protection solutions.
- *New techniques* which identify unknown threats, such as malware that has never been seen before, have become more and more common.

The combination of both above techniques, often referred to as "next-gen", appears as optimal solution od end point security.

The basic features of traditional techniques include:

- *Anti-malware/antivirus* that is based on detection of signature of known malware. Malware engines should have the ability to inspect not just executables but also other code such as malicious JavaScript found on websites.

- *Application lockdown* which Preventing malicious behaviors of applications, like a weaponized Office document that installs another application and runs it.

- *Behavioral monitoring/Host Intrusion Prevention Systems (HIPS)* that represents foundational technology protects computers from unidentified viruses and suspicious behavior. It should include both pre-execution and runtime behavior analysis.

- *Web protection* that provides URL lookup and blocking of known malicious websites. Blocked sites should include those that may run JavaScript to perform cryptomining, and sites that harvest user authentication credentials and other sensitive data (Nyemkova et al., 2020).

- *Web control* that represents endpoint web filtering allows administrators to define which file types a user can download from the internet.

- *Data loss prevention (DLP)* that prevents an adversary to be unnoticed. DLP feature would be able to detect and prevent the last stage of some attacks, when the attacker is attempting to exfiltrate data. This is achieved by monitoring a variety of sensitive data types.

The new features of traditional techniques include:

- *Machine learning* where are multiple types of machine learning methods, including deep learning neural networks, random forest, bayesian, and clustering. The advantage of machine learning is that it can detect malware that has never been seen before, ideally increasing the overall malware detection rate. Organizations should evaluate the detection rate, the false positive rate, and the performance impact of machine learning-based solutions.

- *Anti-exploit technology* is designed to deny attackers to use the tools and techniques they rely on in the attack chain. This technology stops the relatively small collection of techniques used to spread malware and conduct attacks, warding off many zero-day attacks without having seen them previously.

- *Ransomware-specific* represents solutions contain techniques specifically designed to prevent the malicious encryption of data by ransomware attack in destructive wiper attacks that tamper with the master boot record. This feature should be able to remediate any impacted files.

- *Credential theft protection* that is designed to prevent the theft of authentication passwords and hash information from memory, registry, and off the hard disk.

- *Process protection (privilege escalation)*, which is protection that determines whether a process has a privileged authentication token inserted as part of an active adversary attack that steals a authentication token.

- *Process protection (code cave)* that prevents use of techniques such as code cave and AtomBombing often used by adversaries looking to take advantage of the presence of legitimate applications. Adversaries can abuse these calls to get another process to execute their code.

- *Endpoint detection and response (EDR)/root* causes analysis are not focused on preventing attacks, but rather analyzing and responding to previously detected incidents.

- *Incident response/Synchronized Security* when endpoint tools should at a minimum provide insight into what has occurred to help avoid future incidents. Ideally, they would automatically respond to incidents, without a need for analyst intervention, to stop threats from spreading or causing more damage (Beley & Chaplyha, 2017). It is important that incident response tools communicate with endpoint security tools and network security tools.

- *Managed Threat Response (MTR)* provides delivers 24/7 threat hunting, detection and response delivered by a team of experts as a fully managed service. Analysts should be able to respond to potential threats, look for indicators of compromise and provide detailed analysis on events that took place, where, when, how and why.

When organization evaluates endpoint innovation, they should not just prefer for one primary feature. Relying on one dominant feature, even if it is best-in-class, means that the innovation solution is vulnerable to single point of failure. Conversely, a defense-in-depth approach, where there is a collection of multiple strong security layers, will stop a wider range of threats. This is what it often refers to as "the power of the plus" – a combination of foundational techniques, plus machine learning, plus anti-exploit, plus anti-ransomware, plus EDR, plus much more.

2.3. Comparison of available endpoint innovation solutions

Comparing products with different features is hard enough, but comparing their performance in simulated attacks, where an attacker's actions are potentially infinite and unknown, is nearly impossible (Doucek et al., 2018a). Therefore, many organizations choose to rely on third party assessments to aid their buying decisions.

Various evaluations are available to organizations in the selection of endpoint innovation solutions in the form of comparative tables and graphs, as shown in Table 1.

Tab. 1 Test of AV-Comparatives Bussines Security, source: according Intercept X

Supplier	Malvare Protection Rate
Avast, Bitdefender, Panda, Sophos, SparkCognition	99,9%
Cisco, Symantec, Trend Micro	99,8%
K7, McAfee	99,7%
Segride	99,6%
FireEis, Microsoft, ESET	99,5%

3. Evaluating endpoint security innovation solution

One of the possible and simple ways to evaluate the endpoint innovation solution is to answer the vendor to the following 10 questions:

- Does the product rely on foundational techniques, modern techniques, or a combination of both? Which specific features are core to the technology?

- How does the product detect unknown threats? Does it utilize machine learning?

- For products claiming to leverage machine learning, what type of machine learning is used? Where does the training data come from? How long has the model been in production?

- What technology exists to prevent exploit-based and file-less attacks? What antiexploit techniques are leveraged, and what types of attacks can they detect?

- Does the product have technology specifically designed to stop ransomware?

- Does the vendor have third party results validating their approach?

- Does the product have an acceptable level of false positives? If a false positive is detected, how easy is it to reduce its impact?

- What visibility into an attack does the vendor provide, such as root cause analysis?

- Does the product automatically respond to a threat? Can it automatically clean up a threat and respond to an incident?

- What level of effort is involved in the deployment and use of the innovation solution?

3.1. Evaluation of benefits

The endpoint innovation solution should add value to the whole business and an effective solution that will enable:

- *reducing cyber risk* so that reduce attack exposure and significantly increase response time in case of infection (Caravelli & Jones, 2019), (Nyemkova et al., 2019);

- *increasing visibility* with gaining deeper and wider insight into safety of property, it allowing to make informed and accurate decisions (Diogenes & Ozkala2019);

- *Increasing productivity* so that it reduces the impact of computer security on the IT team and users throughout the organization (Doucek et al., (2018b);

- *saving money* by switching from point products to a computer security system what allows to reduce on-board costs, integration and training, as well as day-to-day system management overhead. Reseller consolidation benefits from non-IT functions such as purchasing and legal;

- *displaying security value* what reduce time spent solving everyday problems, a computer security system can free up IT teams to work on business-oriented projects. Increased protection and resulting reduction of user downtime also allows a wider organization to appreciate the value of security (Pavlicek & Střížová, 2017).

4. Conclusion

Now the management of organizations should consider the innovation of protecting the entire digital environment. When solving and implementing endpoint security, management of organization need to know that it is only one part of the overall security strategy. Some specific technologies to consider along with innovation of endpoint protection include full disk encryption, mobile device management, mobile security, secure email gateway, specialized server or virtual machine protection, and Synchronized Security between endpoint and network devices.

From the customer's point of view, it is good for a single supplier to provide innovation solutions that together create a consistent solution protection and enforcement throughout the organization. Cooperation with one the vendor can provide better security, reduce administration, and reduce costs. Cooperation with one the vendor can provide better security, reduce administration, and reduce costs.

5. Acknowledgement

6. References

Beley, A. & Chaplyha, V. (2017). The Application of Neural Networks for the Intelligent Analysis of Multidimensional Data. In Proceedings of 2017 4th International Scientific-practical Conference Problems of Infocommunications-Science and Technology (PIC S&T). Kharkiv, Ukraine. New York: IEEE, pp. 440-404.

Caravelli, J. & Jones, N. (2019) Cyber Security Threads and responses for Government and Bus., Santa Barbara: Praeger.

Danel, R., Kozel, R., Chlopecký, J., Vilamová, Š., & Piecha, M. (2015). Information Support for Sales Management in the Company OKD a. s., 11th International Conference on Strategic Management and its Support by Information Systems, Uherské Hradiště, Czech Republic, pp. 46-54.

Danel, R., & Neustupa, Z. (2016). Information Support for Brownfield Revitalization Projects. In Computer Science and Information Technology – CSIT 2016, Lvov, Ukraine, pp. 111-115.

Doucek,P, Pavlicek, A. & Luc, L. (2018a). Internet of thigs or surveillance of things? 11th IFIP WG 8.9 Working Conference on Research and Practical Issues of Enterprise Information Systems, CONFENIS 2017, Springer.

Doucek, P., Kuncová, M., Nedomová, L. (2018b). The penetration of ICT into the economy - technical infrastructure in the V4 countries. 12th IFIP WG 8.9 Working Conference on Research and Practical Issues of Enterprise Information Systems, CONFENIS, Poznan, Poland: Springer, pp. 69-78.

Diogenes, Y. & Ozkala, E. (2019) Cyber security – attack and Defense Strategies, 2nd edition, Birmingham: Pakt Pub.

Forbes.com (2018), Chief Information Security Officer Priorities For 2019, [Online], Available https://www.forbes.com/sites/oracle/2019/01/17/chief-information-security-officer-priorities-for-2019/#62641d8a6937 [15 May 2019]

Franke, D. (2016) Cyber Security Basics: protect your organization by applying the fundamentals, South Carolina: CreateSpace Independent Publishing Platform.

Chlopecký, J., Moravec, L., Danel, R., Ameir, O. External analysis for the purpose of strategic decision-making of heating company. 13th International Conference on Strategic Management and Its Support by Information Systems, SMSIS 2019. Ostrava, Czech Republic: VSB-Technical University of Ostrava, pp. 33-41.

Maksymovych, V., Nyemkova, E., Shevchuk, M. (2019). Statistic properties and cryptographic resistance of pseudorandom bit sequence generators. Advances in Computer Science for Eng. and Education, Springer.

Nyemkova, E., Vojtusik, S., Ragulin, A., Chaplyha, V. (2019). Real-time authentication of electronic devices in cyber-physical systems. In Proceedings of 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T). Cham: Springer, pp. 589 – 592. Scopus.

Nyemkova, E., Chaplyha, V., Ministr, J., Chaplyha, V. (2020). Advanced Methods and Means of Authentication of Devices. Lecture Notes on Data Engineering and Communications Technologies: Business Information Data-Centric Business and Applications. Cham: Springer, pp. 205-228, ISBN 978-3-030-35648-4.

Pavlicek, A. & Střížová, V. (2017), Cross culture differences on Facebook. 25th Interdisciplinary Information Management Talks: Digitalization in Management, Society and Economy, IDIMT 2017. Verlag, pp. 137-146.

Rozehnal, P. & Danel, R. (2019). Possibilities of ITIL and PCF mapping. 13th International Conference on Strategic Management and Its Support by Information Systems, SMSIS 2019. Ostrava, Czech Republic: VSB-Technical University of Ostrava, pp. 410-417.

Sophos (2018), The Dirty Secrets of Network Firewalls, Sophos, [Online], Available https://www.insight.com/en_US/content-and-resources/brands/sophos/the-dirty-secrets-of-network-firewalls-infographic.html [30 April 2019]

Sophos (2019), 7 uncomfortable truths of Endpoint Security: A Sophos report, [Online], Available https://news.sophos.com/en-us/2019/03/05/7-uncomfortable-truths-of-endpoint-security-a-sophos-report/ [30 January 2020]

KNOWLEDGE TRANSFER IN AGILE DIGITAL ENVIRONMENT: IMPLICATIONS FOR THE COMPANY'S INNOVATION POTENTIAL DEVELOPMENT

Olga Revutska, Anastasiia Mazurchenko

olga.revutska@tul.cz, anastasiia.mazurchenko@tul.cz

Keywords

Knowledge sharing, agile management, innovation, human resources

Abstract

Nowadays, knowledge and intellectual capital are among the most valuable assets necessary for the company's economic growth in the global competition, as well as for advancement of the organizations´ innovative capabilities. Knowledge transfer within the company facilitates problem solving and reducing costs, improves quality of the decision-making processes, employees' productivity and organizational innovativeness. From this point of view, an agile digital environment supports creative thinking, encourages employees´ to experiment and contributes to the individual professional development. The paper aims to provide a review of the literature related to the topic and analyses the impact of agile digital environment on the company´s innovation potential on data collected in Digital Business Global Executive Survey (2017) based on the opinion of 4,300 respondents from all over the world. In this part of the paper data related to adoption of the innovations in the organizations were determined and analysed to demonstrate characteristics of maturing digital companies in accordance with the principles of the agile management. Findings in the paper confirmed that agility and digital environment have positive effect on improving a knowledge sharing within the company and enhancing companies' innovation capabilities.

1. Introduction

The importance of human capital and knowledge for the companies' operation and development has been a subject of many scientific discussions in recent years. The transformation of society characterized by the popularization of information and communication technologies and networks, is becoming essential to achieving companies' labour productivity in a knowledge-based and innovation-oriented economy (Atkinson & Steward, 2012; Klima, 2004). Modern organizations are characterized by a huge potential of data and knowledge. Their sharing and use in conjunction with human factor is an important source of business competitiveness (Rábová, 2006). In view of the fact that the companies' growth depends primarily on the application of the employees' knowledge and intellectual capital (Jermář, 2012), knowledge transfer supports organizational innovativeness (Ho & Kuo, 2013) and leads to better understanding of the customers´ and other stakeholders´ needs (Matošková, 2019). It should also be mentioned that knowledge transfer is vital to exploit core employees´ competencies and achieve a sustainable organisational development (Anwar, 2017).

The current business environment is characterized by rapid technological advances, an emphasis on strong competition and digitalization. In such conditions, a success of business entities is largely

determined by the ability to respond to changes fast (Šochová & Kunce, 2014) and introduce innovations in order to achieve a more favorable position in the market (Rozmajzl & Hovorková, 2012). Traditional approaches based on planning and management of each business activity to achieve a clearly defined goal ceases to meet the needs of knowledge-based economy. The conventional methods are being replaced by agile approaches that support a process of continuous knowledge transfer (Šochová & Kunce, 2014), promote organizational flexibility, and contribute to increasing employee engagement, motivation and willingness to share knowledge within the organization (Swift et al., 2010). The aim of the paper is to introduce the aspect of agile and digital environment in the knowledge transfer, as well as to introduce selected findings about influence of the digitalization on the company's innovation potential.

2. Knowledge transfer as a driver of the companies' innovative activity

According to Pitra & Mohelská (2015, p. 22) *knowledge* can be defined as an intangible phenomenon that allows understanding the nature of the facts, processes and behavior of real objects, successful mastering of various activities and strategy selection for successful overcoming of obstacles to achieving the chosen goal. In the Czech Terminology Database of Library and Information Science, *knowledge* is explained as the ability of human or any other intelligent system to store, communicate and process information into systematic and hierarchical structures. Omotayo (2015) points out that creation and acquisition of a new knowledge take place in a spiral. This process starts from sharing individual internal tacit knowledge by socializing with others or capturing it in digital or analogue form, and continues by internalizing in a form of an *explicit knowledge*. *Tacit knowledge* is those that is impossible or difficult to clearly define in individuals' mind, whereas *explicit knowledge* is defined as sentences, organized data, computer programs and other specific forms (Sezgin & İplik, 2018). A constant knowledge transfer is fundamental for generating new ideas (Amir & Parvar, 2014) and leads to innovations (Omotayo, 2015).

Miklosik et al. (2012) explain a *knowledge transfer* as an effective use and dissemination of created intellectual property. It is worth noting that *knowledge transfer* takes the form of internal and external communication between teams made up of employees of the individual components of the economic system (Pitra & Mohelská, 2015). In such case, the way of communication is more important for knowledge transfer than shared content. Choosing an appropriate communication style encourages the interest and involvement of all actors in active knowledge sharing. Conversely, an inappropriate style of communication discourages many creative individuals from working with other members on innovative solution. As Rábová (2006) clarifies, an effective implementation of the knowledge transfer to a big extend is influenced by the following factors:

- people and culture in an organization,
- business processes that support the entire knowledge life cycle,
- Information and communication technologies.

It should also be empathized that knowledge transfer cannot be forced by the companies' managers but can only be encouraged and facilitated by them (Gibbert & Krause, 2002). Analyst estimates suggest that American companies in the Fortune 500 lose a total of $31.5 billion per year because of the employees failing to share knowledge effectively (Myers, 2015). In such conditions creating an environment where employees are inspired to experiment, providing them with information and communication technologies (Seba et al., 2012) and motivating them by human resource practices (Chen & Huang, 2009) is essential for efficient knowledge transfer.

It seems to be clear that encouraging new ideas is an integral part of good organisational management. A suitable corporate culture that fosters a creative workplace atmosphere, multifunctional

collaboration, openness to change, and aspiration to the professional growth is key to driving innovation activities. According to the Oslo Manual (OECD & Eurostat, 2018) *innovation activities* include all developmental, financial and commercial activities undertaken by a firm that are intended to result in a new or improved product or process (or combination thereof), differs significantly from the unit's previous products or processes, has been made available to potential users or brought into use by the firm.

Innovation activity requires creative thinking, as well as engagement, persuasiveness and the ability to overcome barriers. If the company is not innovation friendly, knowledge transfer is bound to fail (Dalkir, 2005, p. 51). On the contrary, if companies will implement an appropriate innovation policy, they will be able to gain not only a competitive advantage, but also develop an image of the pioneer on the market and make a profit due to low manufacturing costs or higher market prices. The importance of knowledge for the companies' innovative activity was investigated by Jermář in 2010 in 52 manufacturing organizations and knowledge-oriented companies mainly located in Pilsen region of the Czech Republic. As a result, key advantages of applying employees' knowledge in day-to-day business operations were identified (see Figure 1).
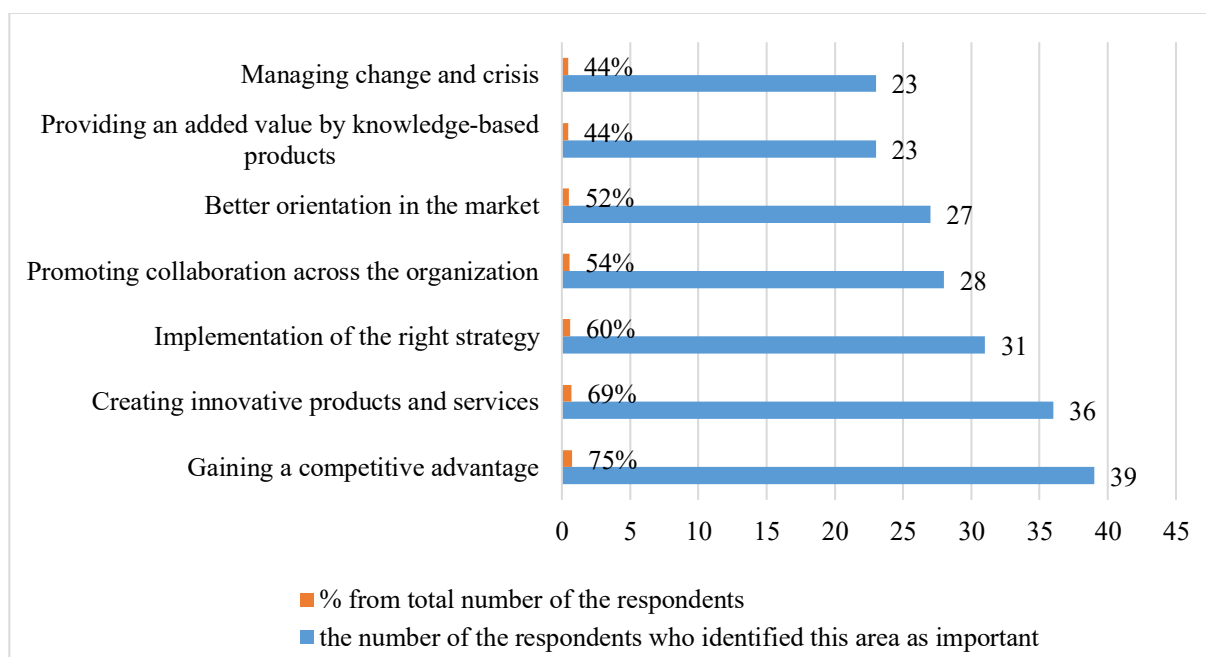
Figure 1. The most significant benefits of working with knowledge within the company; Source: own elaboration based on (Jermář, 2012)

3. The role of agile management in supporting innovations

The success of the companies' innovation activities is directly linked to the human factor. New discoveries, technologies and ideas for their commercial use come primarily from employees. For this reason, one of the main prerequisites for a functioning companies' innovation systems and the economy in a whole is a sufficient number of motivated people equipped with the necessary knowledge and competences (Národní Inovační Strategie České Republiky, 2011, p. 32). Flexibility in personnel management and a creative corporate culture support the employees' inventive activity, which is an essential part of the innovation process. Therefore, a development of innovation activities requires agility of the human resources management. The traditional hierarchy is being replaced by agile management, at which small multifunctional teams have more freedom in decision-making and

maximally oriented on results (Šochová, 2019). The fundamental differences between elements of conventional and agile management are demonstrated in the Table 1.

Table 1. Comparison between the traditional and agile management elements

Traditional management elements	Agile management elements
Processes and tools	Individuals and interactions
Comprehensive documentation	Working software
Contract negotiation	Customer collaboration
Following a plan	Responding to change

Source: own elaboration based on (Manifesto for Agile Software Development, 2001)

Definition in the Otto's dictionary (1888, p. 446) describes the term *agile* as nimble, adept, energetic, quick, active, flexible, and responsive to change. *Agility in human resources management* means working with talents and their personal development based on active collaboration with the team, helping to ensure rapid response to any changes in external environment (Šochová & Kunce, 2014, p. 175). In accordance with the Manifesto for Agile Human Resource Development (2017), the mission of the human resource management is to support the people's and team's capabilities, help them build the environment they need, facilitate and nurture personal growth, to harness employee's different strengths and talents.

Agile methods contribute to a continuous learning environment (Dingsøyr et al., 2012) where knowledge sharing can take place through close interaction between team members (Bannerman et al., 2012) or through frequent interactions and collaboration with the client (Dorairaj et al., 2012). Scrum, as one of the most widespread iterative, incremental and team-based agile methods (Miklosik et al., 2012), can also be considered as an important tool to assist in implementing knowledge management practices in organizations (Tenório et al. 2020).

Review of the literature indicates an emphasis on the strategic importance of knowledge. The synergy effect of employee's knowledge and intelligence creates a competitive advantage and promotes innovation potential of the companies. It also appears that the prerequisites for successful knowledge sharing comply with the principles of agile organizations.

4. Adoption of the innovations in a digital environment: analysis of the secondary data

In the research conducted by MIT Sloan Management Review and Deloitte in 2017, 4,300 employees and senior managers were asked about a readiness to adopt their organizations to work in a digital environment defined as an integrated communication environment, wherein digital technologies and devices are used to improve communication, manage the content, and drive new value generating business models. Individuals from the 123 countries from organisations of various sizes agree that pace of business (23 %), culture and mindset (19 %), flexible workplaces (18 %) and productivity (16 %) are among the biggest differences distinguish work in digital environment from traditional. At the same time, 44 % of the respondents believe that it is needed constantly update their skills in order to stay effective and productive in a digital environment. Figure 2 demonstrates the areas in which employees get support from their organizations to create innovations in the digital environment. However, many organizations (42 %) still do not analyse the reason why the implementation of innovations was unsuccessful, and do not apply lessons learned.
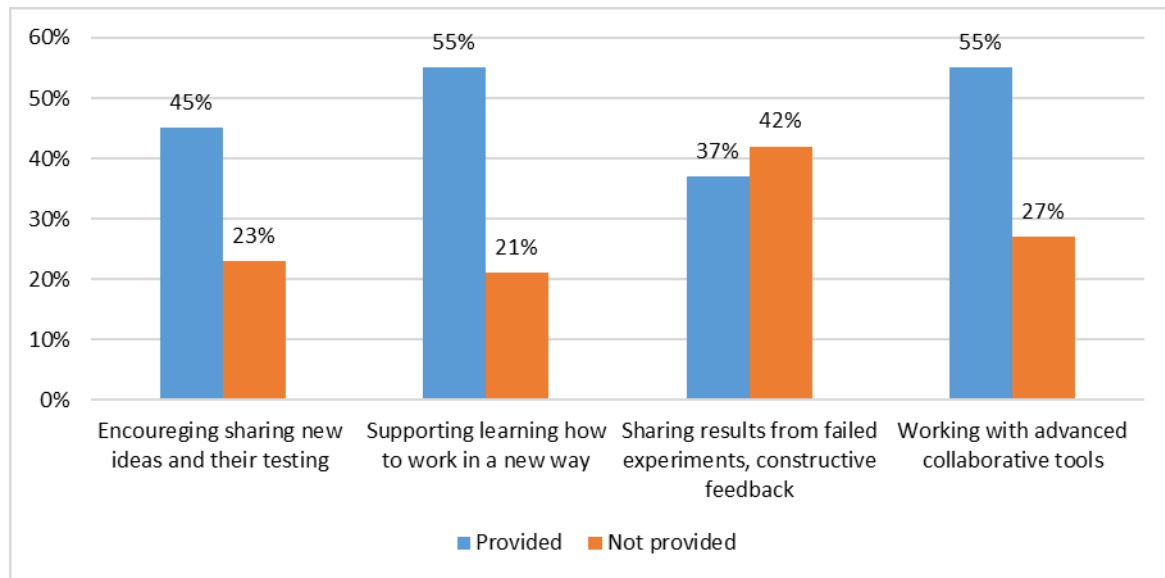
Figure 2. Employees' innovation potential development supported by organization in the digital environment;

Source: own elaboration based on (Kane et al., 2018)

The survey showed that the characteristics of maturing digital companies, i.e. those rated 7-10 on a scale of 1 to 10 against an ideal organization utilizing digital technologies and capabilities to improve processes, engage talent, and drive new value generating business models, correspond to the principles of agile management (see Table 2).

Table 2. Intersection of agile management principles and characteristics of maturing digital companies

Agile management principles	Characteristics of maturing digital companies
Welcome changing requirements	Facilitating change, encouraging new ideas, developing a culture of experimentation
Self-organizing teams	Decision-making at the lower levels of the organization
Cross-functional teams	Enhancing cross-functionally within broader organization
Customer collaboration	Enhancing collaboration with customers
Continuous improvement (efficiency) mindset	Encouraging feedback and new ways of working, promoting of organizational learning

Source: own elaboration based on (Manifesto for Agile Software Development, 2001; Kane et al., 2019)

As a result, the agile digital business environment contributes to creating and developing a culture of experimentation, continuous learning and inventiveness. However, respondents also point to areas where improvements can help the organization better navigate the digital trends: creating the conditions for people to experiment, empowering people to think differently, getting people to collaborate across boundaries, and supporting continuous self-development (Kane et al., 2019).

Modern digital technologies, which include the internet of things, next-generation telecommunication networks, big-data analytics, artificial intelligence using deep learning and blockchain technology, are highly inter-related (centralized) and became a subject of many debates concerning, among other things, business operation in times of pandemic Covid-19 (Ting et al., 2020).

5. Discussion

Despite the fact that knowledge has a strategic importance for the development of innovative business potential, a third of the Czech companies still does not educate their employees and does not even work with the development of knowledge (Urbancová et al., 2016). Extend of awareness about agile methods contributing to knowledge transfer is low and a major part of the Czech companies does not apply them at all (Rozmajzl & Hovorková, 2012). Results from the 2018 Digital Business Global Executive Survey has shown that 32 % of the organizations from 125 countries do not provide enough time, money and people to support innovations (Kane et al., 2019). In addition, a lack of qualification and incompetence of project team members greatly endangering the success on innovative projects (Krátký & Lokaj, 2015). Another research (Louženská et al., 2017) confirmed that personal development, employees' way of thinking and the corporate culture (39 %) are among the biggest barriers blocking innovations.

The willingness of employees to share knowledge also depends on communication style and relationships in organizations (Urbancová et al., 2016). A study conducted by Matošková (2019) confirmed a positive statistically significant correlation between employee motivation and knowledge sharing in the organization, with intrinsic motivation driven by self-realization and self-affirmation being stronger than desire to get monetary rewards and benefits (Jain et al., 2019).

This exploratory research is limited to a literature review. Limitations on secondary data are related to the number of respondents and their geographic (32 % from the USA). The results of the paper are means to stimulate new researches to assess empirical findings in a practical context. The challenge is to carry on an in-depth analysis of the benefits of the agile and digital environment for business operations and development.

6. Conclusion

The paper aimed to provide a review of the literature related to the impact of the agile digital environment on the company´s innovation potential. The findings in the paper confirmed that agile processes welcome changing requirements and transform them into the customer's competitive advantage, stimulating cross-functional collaboration between self-organized teams empowered with the decision-making authority, and enhancing a continuous learning environment and inventive mindset. Digitally maturing companies, on their part, are developing a culture of experimentation, encouraging new ideas to be shared and tested at all levels of the organization, and providing advanced collaborative tools to facilitate better communication. The synergy of both creates a positive effect on the knowledge sharing and companies' innovation potential.

7. Acknowledgement

8. References

Agile. (1888). In Ottův Slovník Naučný. Illustrovaná Enciklopedie Obecných Vědomosti. První díl. Praha: J. Otto. Retrieved from https://pdfknihy.maxzone.eu/books/OSN/otto01.pdf

Amir, R.,& Parvar, J. (2014). Harnessing Knowledge Management to Improve Organisational Performance. International Journal of Trade, Economics and Finance, 5(1), 31-38. DOI: 10.7763/IJTEF.2014.V5.336

Anwar, C. M. (2017). Linkages Between Personality and Knowledge Sharing Behavior in Workplace: Mediating Role of Affective States. E&M Economics and Management, 20(2), 102-115. DOI: 10.15240/tul/001/2017-2-008

Atkinson, R. D., & Stewart, L. A. (2012). The 2012 State New Economy Index: Benchmarking Economic Transformation in the States. Washington: The Information Technology & Innovation Foundation. Retrieved from http://www2.itif.org/2012-state-new-economy-index.pdf

Bannerman, P. L., Hossain, E., & Jeffery, R. (2012). Scrum Practice Mitigation of Global Software Development Coordination Challenges: A Distinctive Advantage? In Proceedings of the 45th Hawaii International Conference on System Sciences (pp. 5309-5318). New York, USA: Institute of Electrical and Electronics Engineers. DOI: 10.1109/HICSS.2012.512

Chen, C.-J., & Huang, J.-W. (2009). Strategic Human Resource Practices and Innovation Performance — The Mediating Role of Knowledge Management Capacity. Journal of Business Research, 62(1), 104-114. DOI: 10.1016/j.jbusres.2007.11.016

Dalkir, K. (2005). Knowledge in Theory and Practice. Burlington, USA: Elsevier Butterworth–Heinemann.

Dingsøyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A Decade of Agile Methodologies: Towards Explaining Agile Software Development. Journal of Systems and Software, 85(6), 1213-1221. DOI: 10.1016/j.jss.2012.02.033

Dorairaj, S., Noble, J., & Malik, P. (2012). Knowledge Management in Distributed Agile Software Development. In Proceedings of the Agile Development Conference 2012. (pp. 64-73). New York, USA: Institute of Electrical and Electronics Engineers. DOI: 10.1109/Agile.2012.17

Gibbert, M., & Krause, H. (2002). Practice Exchange in a Best Practice Marketplace. In T. H. Davenport and G. J. B. Probst (Eds.), Knowledge Management Case Book: Siemens Best Practices. (pp.89-105). Erlangen, Germany: Publicis Corporate Publishing.

Ho, L.-A., & Kuo, T.-H. (2013). How System Quality and Incentive Affect Knowledge Sharing. Industrial Management & Data Systems, 113(7), 1048-1063. DOI: 10.1108/IMDS-01-2013-0015

Jain, A.,Gupta, B., & Bindal, M. (2019). A Study of Employee Motivation in Organization. International Journal of Engineering and Management Research, 9(6), 65-68. DOI: 10.31033/ijemr.9.6.11

Jermář, M. (2012. Rozvoj Znalostního Potenciálu Firem - Inspirace pro Management Lidských Zdrojů. E&M Economics and Management, 15(2), 85-93.

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2019). Accelerating Digital Innovation Inside and Out: Agile Teams, Ecosystems and Ethics. MIT Sloan Management Review and Deloitte University Press. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/deloitte-digital/lu-accelerating-digital-innovation.pdf

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2018). Coming in Age Digitally: Learning, Leadership and Legacy. MIT Sloan Management Review and Deloitte University Press. Retrieved from https://assets.ctfassets.net/zw48pl1isxmc/Bs3y90xB61smjIfmRNVST/9a3a828e6d653a8a51a1cf4c7ffc5ae8/Optimizely_CoAD__1_.pdf

Klima, E. (2004). Information Society in the European Union – History of Common (?) Policy. Space-Society-Economy, 6, 1-10.

Knowledge. (n. d.). In The Czech Terminology Database of Library and Information Science. Retrieved from https://aleph.nkp.cz/F/TMT7G63CEMDC5Y4BF5Q8UN9136MRKAU8I9BRAL4RRPMMSNB95M-52709?func=full-set-set&set_number=114620&set_entry=000006&format=999

Krátký, J., & Lokaj, J. (2015). Projektové Řízení v ČR 2015. Zpráva o Výsledcích Šetření. PM Consulting and Společnost pro Projektové Řízení Česká republika: Brno, Czech Republic. Retrieved from https://www.ipma.cz/media/1282/2015_projektové-řízení-v-čr-vysledky.pdf

Louženská, P., Pitter, L., Sobotka, A., Libová, A., Tomanová, D., Jílková, V., & Freiwilligová, M. (2017). SPINN. Studie Problémů s Inovacemi. 2Fresh. Retrieved from https://assets.website-files.com/5aeaf0ae05cf342ea77798ae/5b2cae34a49e8ec216050d72_Spinn%20study.pdf

Manifesto for Agile HR Development. (2017). Retrieved from https://www.agilehrmanifesto.org/

Manifesto for Agile Software Development. (2001). Retrieved from https://agilemanifesto.org/

Matošková, J. (2019). Why Employees Share Their Knowledge. E&M Economics and Management, 22(2), 83-96. DOI: 10.15240/tul/001/2019-2-006

Miklosik, A., Hvizdová, E., & Žák, Š. (2012). Scrum as the Method Supporting the Implementation of Knowledge Management in an Organization. International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 6(3), 372-375.

Myers, C. G. (2015). Is Your Company Encouraging Employees to Share What They Know? Harvard Bus. Review. Retrieved from https://hbr.org/2015/11/is-your-company-encouraging-employees-to-share-what-they-know

Národní Inovační Strategie České Republiky. (2011). Ministry of Industry and trade: Prague. Retrieved from https://www.mpo.cz/assets/dokumenty/44911/50470/582602/priloha002.pdf

OECD & Eurostat. (2018). Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation. 4th ed. , The Measurement of Scientific, Technological and Innovation Activities. OECD Publishing: Paris/Eurostat, Luxembourg. DOI: doi.org/10.1787/9789264304604-enOECD

Omotayo, F. O. (2015). Knowledge Management as an important tool in organizational management: a review of Literature. Library Philosophy and Practice, 1238, 1-23. Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=3330&context=libphilprac

Pitra, Z., & Mohelská, H. (2015). Management Transferu Znalostí: Od Prvního Nápadu ke Komerčně Úspěšné Inovaci. Praha, Česka Republika: Professional Publishing.

Rábová, I. (2006). Objektové Modelování Podnikových Znalostí a Znalostní Management. In Proceedings of the International Scientific Days 2006 "Competitiveness in the EU - Challenge for the V4 Countries". (pp. 1578 - 1583). Nitra, Slovakia: Slovak Agricultural University in Nitra. Retrieved from https://spu.fem.uniag.sk/mvd2006/zbornik/sekcia8/s8_rabova_ivana_184.pdf

Rozmajzl, P., & Hovorková, I. (2012). Agilní přístupy v projektovém řízení. In Proceedings of the Liberec Economic Forum 2012 (pp. 93-98). Liberec, Czech Republic: Technical University of Liberec.

Seba, I., Rowley, J., & Lambert, S. (2012). Factors Affecting Attitudes and Intentions Towards Knowledge Sharing in the Dubai Police Force. International Journal of Information Management, 32(4), 372-380. DOI:10.1016/j.ijinfomgt.2011.12.003

Sezgin, A. A., & İplik, E. (2018). From Personal Knowledge Management to Corporate Knowledge Management. In F. Di Virgilio (Ed.), Social Media for Knowledge Management Applications in Modern Organizations. (pp. 169-189). Hershey, USA: IGI Global. DOI: 10.4018/978-1-5225-2897-5.ch008

Swift, M., Balkin, D. B., & Matusik, S. F. (2010). Goal Orientations and the Motivation to Share Knowledge. Journal of Knowledge Management, 14(3), 378-393. DOI: 10.1108/13673271011050111

Šochová, Z. (2019). Agilní HR: Jak Změnit Human Resources v Agilní Organizaci. Retrieved from https://sochova.cz/novinky/novinky/agilni-hr-jak-zmenit-human-resources-v-agilni-organizaci.htm

Šochová, Z., & Kunce, E. (2014). Agilní Metody Řízení Projektů. 1st ed. Brno: Computer Press.

Tenório, N., Pinto, D., Silva, M. J., Carnevale de Almeida, A., & Bortolozzi, F. (2020). Knowledge Management in the Software Industry: How Scrum Activities Support a Knowledge Management Cycle. Navus: Revista de Gestão e Tecnologia, 10, 1-13. DOI: 10.22279/navus.2020.v10.p01-13.928

Ting D. S. W., Carin L., Dzau V. & Wong T. Y. (2020). Digital technology and COVID-19. Nature Medicine. 26. DOI: 10.1038/s41591-020-0824-5

Urbancová, H., Vnoučková, L., & Laboutková, S. (2016). Knowledge Transfer in a Knowledge-Based Economy. E&M Economics and Management, 19(2), 73-86. DOI: 10.15240/tul/001/2016-2-005

SOCIAL AND TECHNICAL ASPECTS
OF RE-DECENTRALIZED WEB

Stanislav Vojíř, Zdeněk Smutný, Jan Kučera

Faculty of Informatics and Statistics
University of Economics, Prague
stanislav.vojir@vse.cz, zdenek.smutny@vse.cz, jan.kucera@vse.cz

Keywords

Blockchain, Cloud Feudalism, Crypto Technology, Decentralized Technology, Information Society, Internet Studies, Platform Capitalism.

Abstract

Although the World Wide Web (or web) is a decentralized network of interconnected documents, highly centralized ecosystem dominated by a few supranational companies has developed on top of these foundations. These companies create their own platforms and offer web-powered services that are used by people and organizations around the world. Many of these platforms or services are free of charge, but the providers often process personal and other users' data not just to improve their services but also for other commercial purposes. The users' data also represent a significant barrier to entry of new companies to the market, because collecting the users' data would require significant investments from the new entrants. At the same time, users often face difficulties when switching service providers because their data are in many cases not easily transferable between services. As a result, these supranational companies have become centers of significant power and influence. This article presents the so-called re-decentralization of the web as a complementary alternative to this present state. The goal of the re-decentralized web is to return to the original ideas of the web as an open space, where everyone can freely publish their content and decide on the use of their data. This paper offers a probe into this issue from a social and technical point of view based on an integrative literature review. In this context, some relevant phenomena such as cloud feudalism, platform capitalism or dictatorship of algorithms were highlighted and referred to. In conclusion, the paper summarizes the current situation, presents challenges for further research, and discusses the possible implications of wider adoption of the technologies based on the idea of a re-decentralized web in the near future.

1. Introduction

Although the World Wide Web (WWW, web) represents a decentralized network from a technological point of view, "*billions of people experience the World Wide Web through a small handful of huge companies*" (WWW Foundation, 2018). A large number of services provided on the Internet or the web are provided by large technology companies such as Google, Facebook or Amazon. Development of the web from its inception to today's concentrated organization has its merits, both in terms of used business models (Barnett, 2018) and in practical terms where the services of large technology companies such as web search engines (e.g. Google or Bing) make it easier for users to use the web (Barabas et al., 2017).

Because a significant number of users access services and contents on the web via a small number of large companies, these companies gain considerable power that can potentially be abused (Barabas et al., 2017). In response to this, but also to other problems, such as the spread of fake news, or the unavailability of the web for a large population of the Earth, the Contract for the Web initiative has been established (WWW Foundation, 2020). This initiative aims at ensuring the web is a safe place that connects people and where knowledge can be shared freely. The so-called re-decentralization of the web also falls within the context of the responses to the current situation on the web. Re-decentralized web can be understood as an effort to return to the original ideas of the web as an open space, where everyone can freely publish his or her content and decide on the use of his or her data (Verborgh, 2020).

The aim of this paper is to present the current perspectives on the issue of re-decentralized web related to social and technical aspects (or issues) and to show examples of activities and applications responding to the consolidation of the web in the hands of a few large supranational companies. Because people's ability to make decisions about how their personal data are handled is considered an important aspect of the re-decentralized web, the authors do not present just the activities that represent a distributed system from the technological point of view. The authors are in favour of a broader understanding of these activities and applications, and thus they also include applications that help people decide what data is collected about them and how they are used by the service providers.

This paper is structured as follows. The following chapter describes the methodology. The next two chapters are focussed on the social and technical aspects of the re-decentralized web. The last chapter discusses the main conclusions.

2. Methodology

In order to achieve the aim of this paper, an integrative literature review was conducted. This approach is particularly useful for new or emerging topics such as the re-decentralized web (Torraco, 2005). In such a case, it is necessary to define the subject and framework of this search. The subject is defined by the topic of this paper focused on social and technical aspects of the re-decentralized web (see Introduction). The following keywords were chosen for search in databases: re-decentralized web, decentralized web, re-decentralization decentralized web, decentralization, decentralized technology, privacy, cryptography, social, society. These keywords were further expanded based on other identified phenomena (e.g. platform capitalism).

The research framework is limited by the databases or other sources used to find relevant articles. The primary sources of literature research were citation and full-text databases Web of Science, Scopus, EBSCO, ACM Digital Library, IEEE Xplore Digital Library and Google Scholar. According to their relevance, several dozens of articles were selected from the search results and analysed in detail. In this exploration analysis, the articles were manually clustered by their topics. The main representants of these clusters can be found in References. The authors primarily focused on review articles and books relevant to the discussed phenomena in their search.

3. Social aspects of Re-Decentralized Web

The Internet in its physical form is a decentralized system of interconnected computers, but this does not apply to the status of subjects that operate via the Internet. An example could be the web where a few search engines of supranational companies play an important role (Machado et al., 2019). These search engines can be seen as centres of power (see Bratton, 2016) because they offer content from

other webpages. Search engines act as a gateway to the consumption of the web content, which gives them considerable power, because the position of a webpage in search results greatly affects the likelihood of a user visiting it. For example, Ledford (2008) states that most people visit only the first three pages of search results and most visits receive webpages that are presented in the first half of the first page of search results. Furthermore, this centralization of power is evident in social networking services or social networks (e.g. Facebook). Algorithms embedded in the social networks often determine what media content will be shown on the user's home page, and this will greatly affect what content will or will not be consumed by the user. Some authors call this a dictatorship of algorithms that offer "personalized" content (Hofstetter, 2014; Gianluigi, 2015). But the centralization of power can also be found in the portfolio of interrelated products and services offered by providers such as Amazon, Apple or Google.

Typically, centralization is crucial to the success of a given social network because the social network must reach a critical mass of users. At the same time, network effect is a barrier to competitions. Barabas et al. (2017) point out that people are starting to use a social network because their friends are already using it. It is not easy for users to switch from one social network to another, as they would also have to convince their friends to follow them to another social network.

Companies operating on the web (e.g. e-shops) collect data about their users (personal data, behavioral data etc.). On the one hand, the data is used to offer more relevant services and content to the users, on the other hand, the data ownership is a basis of business models of these companies (Fader, 2012). Usually, the users of these applications cannot easily and freely use the data about themselves. However, barriers to data usage are not necessarily technical or legal (e.g. GDPR). As mentioned above, they may be associated with the inability of the users to easily switch the service providers due to the network effects or the employer requirements. In this context, some also talk about the so-called cloud feudalism (Bratton, 2016) or platform capitalism (Srnicek, 2016), where the platform means a set of interconnected services offered by large companies (e.g. Amazon, Apple, or Google). These perspectives indicate that a user (human or organization) is locked in the particular platform because once data (e.g. documents, databases) is provided to some platform, it might be difficult to transfer the data and use in another service outside of the platform.

Barabas et al. (2017) state that the consolidation of entities providing web services is not a negative phenomenon in itself, as this consolidation has made the web easy to use by users. Problematic is the consolidation of power in the hands of a relatively small number of subjects, which largely determine what users see on the web. An example is the scandal of influencing the US presidential elections in 2016 with a campaign led with the contribution of the company Cambridge Analytica (Wong, 2019; Cadwalladr & Graham-Harrison, 2018).

Especially in the last 20 years this new arrangement of power has been built with the development of the so-called Web 2.0 (Bratton, 2016). Along with the acceptance and adoption of Internet-based technologies at the personal, organizational and social levels, new sociotechnical structures have emerged beyond the territorial sovereignty of individual states (see Bratton, 2016). Due to their importance, large multinational technology companies sometimes find themselves in a situation where their activities are of interest to states. There have been also first conflicts. E.g. in 2010, Google discontinued offering of Internet search in China, allegedly due to cyber-attacks on Google and denial of search censorship (Tan & Tan, 2012). However, it should be added, that Google later worked on a censored search engine for China, but the development was halted after information on its development leaked to the public on 1st August 2018 (Charbonneau, 2019; Tiku, 2019).

This development towards to the centralization of power is currently criticized in several academic and scientific fields such as sociology, political science, economics, media studies or philosophy (Colás, 2019; Bratton, 2016; Srinivasan, 2017; Srnicek, 2016). The current centralization of the web poses a danger, for example, in the form of censorship risk or social exclusion as a result of preventing

access to the service by its provider (Barabas et al., 2017), violations of democratic principles, favoring certain content, loss of market diversification, or a gradual transformation of the understanding of privacy among younger generations.

Efforts to offer decentralized solutions that take away power from the large platforms are supported especially by the young generation. An example might be the young Japanese who started switching to the decentralized Mastodon service due to censorship on Twitter (Verborgh, 2020). It should be added that this alternative in the form of the re-decentralized web will only complement the current offering of the large companies. Nevertheless, it can be assumed that the share of these services may increase significantly in the future.

In particular, the younger generations, known as digital natives (see Prensky, 2001), are aware (Andrade & Monteleone, 2013), that their right to use personal data is limited in favor of applications (most commonly multinational platforms). On the other hand, from a web application provider's point of view, this is a kind of indirect payment for using its service. A disadvantage of this established business model may be, for example, low transparency of further handling of the user data outside the scope of the application (e.g. third parties).

Currently, we can observe emergence of new business models that try to change this paradigm for the benefit of users by emphasizing better control and handling of the personal data. The aim is to prioritize privacy over personalizing services that are "free" and to strengthen democratic principles over censorship. An example is the Brave browser. Its users are rewarded for watching of adverts with a crypto token BAT, which they can then use to support their favorite portal or service (Brave, 2020). For users of decentralized solutions can be also important, that cryptography can be used to ensure that even organizations involved in the operation of decentralized applications or services do not have access to user's data (see, for example, Nilizadeh et al., 2012).

The social aspects of introducing the re-decentralized web principle as a new relationship between the users and the service providers do not only concern the younger generations (Alpha, Z, Y), but also the older generations (X, Baby boomers). The question is whether the older generation users will be willing to accept these changes (Taipale, 2019).

4. Technical aspects of Re-Decentralized Web

If we look at the topic of re-decentralized web from the technical point of view, two basic levels of decentralization can be distinguished: the level of users and the level of creation and operation of truly decentralized applications.

From the users' point of view, it is usually a requirement to ensure privacy – to prevent users' tracking and to ensure anonymity. From a relatively mild attitude, including preventing personalization of ads, to the ideas of Crypto Anarchy (May, 1996). These ideas are not new, but they are gaining attention periodically and in recent years have been supported by technologies that allow the user to be authenticated despite maintaining his or her relative anonymity (Goldenfein & Hunter, 2018; Patel et al., 2019). There are technological tools to ensure this degree of decentralization, for example, by blocking the tracking of user behavior on the web. A basic level of such protection is now provided even in standard web browsers. For example, Firefox turned on enhanced tracking protection in autumn 2019 (Wood, 2019). As far as the anonymization of behavior itself is concerned, the encryption of communication and the support of anonymous payments for content also play a significant role. For the possibility of anonymous payments, cryptocurrencies become a phenomenon (see Ban et al., 2019), e.g. Bitcoin (Nakamoto, 2008) and its derived variants, Ethereum and others. In the field of hiding users' network communication, the best-known variant is Tor Project (2020). Although recently researches appeared showing that the anonymity of Tor users may not be fully

guaranteed due to the possibility to reconstruct the private encryption keys (Kadianakis et al., 2018). It can be said that, with conventional tools, a certain degree of anonymity can be achieved and the potential responsibility for behavior monitored by state institutions can be dispensed (Jordan, 2019; Zhuo et al., 2018; van Hardeveld et al., 2017).

From the technological point of view, the issue of decentralization of the applications themselves and the data processed by them seems to be more interesting. There are requirements for ensuring user authentication, data storage and possibly decentralization of applications, too. In meeting these requirements, we can admit that we are returning to the concept of a decentralized and free Internet.

Looking at a conventional web application or web service, the application has one provider (owner). The users use the functionality of the web application, their data are stored on the servers of the provider and the provider has a full control over the offered functionality and services. The provider has the user data at its disposal and it can use the data for analyzes, predictions etc. Applications and services can be interconnected, but in this case, the usage of e.g. authentication using some social network account strengthens the position of the social network itself and provides the provider additional data about the behavior of the user, even the behavior related to other web pages and services. Centralized services are simpler to implement and monetize and it is easier to predict the users' behavior.

Compared to the conventional applications, decentralized applications provide the users with more control over their data (see also Halpin, 2019). The data is not stored in the data center of the service owner, but the user can choose from more data storage providers or the user can store the data on his or her own server. In case of change of the service terms or conditions, the user may move the data elsewhere. User authentication is usually provided by the community (e.g. using blockchain technologies). Similar to the data storage, there is no dependence on a specific service owner, and thus no one can simply block the user or require to check his or her real identity. The disadvantage of using the concept of decentralization is that data interconnection and user-friendly implementation are more demanding. However, there are working solutions that will be addressed in the following paragraphs.

To authenticate users, blockchain-based technologies can be used, which are similarly used for cryptocurrency transactions. These trends can be found in in the field of authentication of users of web applications and also as in the Internet of Things (Durand et al., 2017). A subject of the current research is the development of standardized protocols to ensure authentication – for example, the DAuth protocol based on Ethereum (Patel et al., 2019).

Decentralized data storage and especially sharing of data among users is still problematic. As we already mentioned, users usually want to share their data in a way that would allow their friends to easily look up and retrieve the data. Semantic web technologies – RDF and Linked Data – could be used to implement these requirements and, in fact, these technologies have already been utilized by some of the decentralized social networks discussed below.

Decentralized social networks represent prominent examples of the decentralized web applications and Diaspora (Paul et al., 2011; Diaspora, 2020) is probably the most popular of such social networks. Diaspora users fully control their data and they can choose a server (called "pod") to store their data. Besides sharing content, users can contribute the computing resources, i.e. they can host a pod. Although Diaspora has been able to build a global availability as well as to attract attention of the research community (Bielenberg et al., 2012; Yeung et al., 2009), its popularity lacks behind its centralized counterparts. Another example of a decentralized social network is Mastodon, a decentralized alternative to Twitter (Mastodon, n.d.).

Solid (2020; Mansour et al., 2016) could be named as an example of a framework for developing decentralized applications. Social networks, blockchain-based identification or webID (Faísca et al., 2016; Wild et al., 2013) could be used to develop the social identity and the author's profile.

Technologies enabling decentralization are not limited to the social networks, but they are emerging in other domains, too. For instance, there are several initiatives aiming at decentralizing the web content itself. Distributed web pages and applications could be built upon the ZeroNet (2020) network which utilizes the BitTorrent network and Bitcoin Wallet as an authentication solution. As an alternative, blockchain or Gaia Storage of the Blockstack decentralized computing network (Blockstack, 2020; Ali et al., 2016) could be used for naming and data storage. Another notable technology platform under active development is the Safe Network (2020).

Re-decentralized web also brings new technological challenges. E.g. (de)centralized search in data on the decentralized web (Trattner et al., 2012; Parreira et al., 2006), knowledge management (Skaf-Molli, 2017; Ayoade et al., 2018) or metadata for linking data sources (Ibanez et al., 2017). There are also initiatives aimed at redesigning the web protocols to better suit the decentralized applications, e.g. the Dat (2020) protocol or Handshake (n.d.), an experimental peer-to-peer naming system and certificate authority.

5. Discussion and conclusion

Social and technical aspects of the re-decentralized web are intertwined. Applications and activities on the technical level provide the users with tools to protect their privacy on the web and to effectively control who and how can collect and use their data. This might weaken the position of the dominant technology companies that apply the advertising-based business models to provide their services free of charge (Barnett, 2018). Because the data collected from the users play a significant role in matching of users to ads and because it is difficult to replicate the databases of the user data, this data represent a barrier to entry for the potential new competitors (Barnett, 2018). Technologies allowing decoupling of the user data from services (see Verborgh, 2020) could make it easier to overcome this barrier. This might attract new service providers which could in turn result in less centralized web on the social level.

While in the 1980s and 1990s was integration a big topic and trend in the area of information and communication technology at the organizational level (e.g. monolithic ERP systems), currently, the trend is decentralization so that companies can react more flexibly and quickly switch between different solutions. Technologies supporting the re-decentralization of the web can also represent enablers of new business models. They can inspire new business models or make them feasible in terms of cost and user experience, as in the case of the Brave browser and its BAT crypto token.

However, availability of the technologies supporting the re-decentralization on the technical level seems like a necessary but insufficient condition for the re-decentralization of the web on the social level. Barabas et al. (2017) point out that the web had been designed as a decentralized system from the start, yet it did not prevent the consolidation around the big international technology companies. According to Raman et al. (2019) centralization is also apparent in Mastodon which, from the technical perspective, has been designed as a distributed system, too. Nevertheless, "*10% of instances host almost half of the users, and certain categories exhibit remarkable reliance on a small set of instances*" (Raman et al., 2019). Re-decentralization of the web on the social level would probably require users and other web participants to change their behaviour.

In order to better understand the phenomena of the re-decentralized web as well as acceptance of the related technologies by society, behavioural and social research aimed at characterizing the behaviour

and preferences of the web users (across generations) should be conducted in the following years. More specifically, such a research could deal with the following topics:

- Web 2.0 usage patterns.

- Attitudes of the web users towards consumption of their personal data by the web applications which is a part of the currently widespread monetization models applied by the application providers.

- Awareness and acceptance of the re-decentralised web.

- Willingness to adapt to the paradigm of the re-decentralised web, most importantly the end users getting truly in control of who, how and when uses their data.

Results of the proposed research should allow us to understand the users' attitudes (e.g. expectations or fears) and it should also indicate the possible ways to support acceptance of the technology in the specific domains. The research should also focus on the inter- and intragenerational differences in order to allow focused supportive actions aimed at the needs of the particular generation or at minimizing the potential negative effects. The proposed research topics might be relevant to researchers in several research fields such as social informatics, information society studies, science and technology studies, internet studies or human-computer interaction.

A shift towards the re-decentralized web on the social level, even if limited, might significantly impact markets which would require organizations to innovate their strategies, business models, the ways they manage users' data or how they promote their products and services (Verborgh et al., 2019). With the growing interest in the decentralized technologies such as blockchain we can expect the interest in the re-decentralized web and the related social and technical aspects to grow as well. Users have become active content creators and they are becoming owners of the data about and related to themselves. Organizations are currently lacking information about the possible impacts of the re-decentralized web (e.g. impacts on the users' behaviour or business models) because the empirical studies dealing with this sociotechnical issue are scarce.

To sum up, the newly proposed technologies and the principles embedded in those technologies affect the behaviour of people which in turn affects the strategies and management of organizations. Re-decentralization of the web can reshape or newly form the sociotechnical structures of the Internet as we know them. Adoption scenarios should be developed in order to help both individuals and organizations accommodate to the changing environment. Understanding the change at the social level is thus a key prerequisite for developing such adoption scenarios.

6. Acknowledgement

7. References

Ali, M., Nelson, J. C., Shea, R., & Freedman, M. J. (2016). Blockstack: A global naming and storage system secured by blockchains. In Proceedings of the Usenix Annual Technical Conference (pp. 181–194). USENIX.

Andrade, N. N. G., & Monteleone, S. (2013). Digital Natives and the Metamorphosis of the European Information Society. The Emerging Behavioral Trends Regarding Privacy and Their Legal Implications. In S. Gutwirth et al., (Eds.), European Data Protection: Coming of Age (pp. 119–144). Springer. https://doi.org/10.1007/978-94-007-5170-5_5

Ayoade, G., Karande, V., Khan, L., & Hamlen, K. (2018). Decentralized IoT Data Management Using BlockChain and Trusted Execution Environment. In Proceedings of the IEEE International Conference on Information Reuse and Integration. IEEE. https://doi.org/10.1109/iri.2018.00011

Ban, Z., Lansky, J., Mildeova, S., & Tesar, P. (2019). A Probe Survey of Bitcoin Transactions Through Analysis of Advertising in an On-Line Discussion Forum. Acta Informatica Pragensia, 8(2), 112–131. https://doi.org/10.18267/j.aip.127

Barabas, C., Narula, N., & Zuckerman, E. (2017). Defending Internet Freedom through Decentralization: Back to the Future?. The Center for Civic Media & the Digital Currency Initiative, MIT Media Lab. https://dci.mit.edu/s/decentralized_web.pdf

Barnett, J. M. (2018). The costs of free: commoditization, bundling and concentration. Journal of Institutional Economics, 14(6), 1097–1120. https://doi.org/10.1017/S1744137418000012

Bielenberg, A., Helm, L., Gentilucci, A., Stefanescu, D., & Honggang Zhang. (2012). The growth of Diaspora – A decentralized online social network in the wild. In Proceedings of the IEEE INFOCOM Workshops. IEEE. https://doi.org/10.1109/infcomw.2012.6193476

Blockstack. (2020). Blockstack: Software for a user owned internet. https://blockstack.org/

Bratton, B. H. (2016). The Stack: On Software and Sovereignty. MIT Press.

Brave. (2020). You deserve a better Internet. https://brave.com/

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Charbonneau, M. (2019). Google executive confirms termination of china search engine project. SNL Kagan Media & Communications Report.

Colás, A. (2019). Internationalism Under Platform Capitalism: Brexit and the Organisation of UK Fast Food Workers. Political Quarterly, 90(4), 620–628. https://doi.org/10.1111/1467-923X.12775

Dat. (2020). Supporting the adoption and development of The Dat Protocol. https://dat.foundation/

Diaspora. (2020). Diaspora. https://diasporafoundation.org/

Durand, A., Gremaud, P., & Pasquier, J. (2017). Decentralized web of trust and authentication for the internet of things. In Proceedings of the Seventh International Conference on the Internet of Things (No. 27). ACM. https://doi.org/10.1145/3131542.3140263

Fader, P. (2012). Customer Centricity: Focus on the Right Customers for Strategic Advantage. Wharton School Press.

Faísca, J. G., & Rogado, J. Q. (2016). Decentralized Semantic Identity. In Proceedings of the 12th International Conference on Semantic Systems (pp. 177–180). ACM. https://doi.org/10.1145/2993318.2993348

Gianluigi, F. (2015). Freedom, Authority and Knowledge on Line: The Dictatorship of the Algorithm. Revista Internacional de Pensamiento Politico, 10, 395–410.

Goldenfein, J., & Hunter, D. (2018). Blockchains, Orphan Works, and the Public Domain. Columbia Journal of Law & the Arts, 41(1), 1–43. https://doi.org/10.7916/jla.v41i1.2037

Halpin, H. (2019). Decentralizing the Social Web. In S. Bodrunova et al., (Eds.), Internet Science: INSCI 2018 International Workshops (pp. 187–202). Springer.

Handshake (n.d.). Handshake. https://handshake.org/

Hofstetter, Y. (2014). Sie wissen alles: Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen. C. Bertelsmann Verlag.

Ibanez, L.-D., Simperl, E., Gandon, F., & Story, H. (2017). Redecentralizing the Web with Distributed Ledgers. IEEE Intelligent Systems, 32(1), 92–95. https://doi.org/10.1109/mis.2017.18

Jordan, T. (2019). Does online anonymity undermine the sense of personal responsibility?. Media, Culture & Society, 41(4), 572–577. https://doi.org/10.1177/0163443719842073

Kadianakis, G., Roberts, C. V., Roberts, L. M., & Winter, P. (2018). "Major Key Alert!" Anomalous Keys in Tor Relays. In S. Meiklejohn, & K. Sako (Eds.), Financial Cryptography and Data Security (pp. 3–19). Springer https://doi.org/10.1007/978-3-662-58387-6_1

Ledford, J. L. (2008). SEO Search Engine Optimization Bible (1.). John Wiley & Sons.

Machado, L. M. O., Souza, R. R., & Graça Simões, M. (2019). Semantic Web or Web of Data? A Diachronic Study (1999 to 2017) of the Publications of Tim Berners-Lee and the World Wide Web Consortium. Journal of the Association for Information Science and Technology, 70(7), 701–714. https://doi.org/10.1002/asi.24111

Mansour, E., Sambra, A. V., Hawke, S., Zereba, M., Capadisli, S., Ghanem, A., Aboulnaga, A., & Berners-Lee, T. (2016). A Demonstration of the Solid Platform for Social Web Applications. In Proceedings of the 25th International Conference Companion on World Wide Web (223–226). ACM. https://doi.org/10.1145/2872518.2890529

Mastodon (n.d.). Social networking, back in your hands. https://joinmastodon.org/

May, T. C. (1996). A Crypto Anarchist Manifesto. In P. Ludlow (Ed.), High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace (pp. 237–240). MIT Press.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf

Nilizadeh, S., Jahid, S., Mittal, P., Borisov, N., & Kapadia, A. (2012). Cachet: A decentralized architecture for privacy preserving social networking with caching. In Proceedings of the 8th international conference on Emerging networking experiments and technologies (pp. 337–348). ACM. https://doi.org/10.1145/2413176.2413215

Parreira, J. X., Debora, Donato, D., Sebastian, M., & Gerhard, W. (2006). Efficient and decentralized PageRank approximation in a peer-to-peer web search network. In Proceedings of the 32nd international conference on Very large data bases (pp. 415–426). ACM.

Patel, S., Sahoo, A., Mohanta, B. K., Panda, S. S., & Jena, D. (2019). DAuth: A Decentralized Web Authentication System using Ethereum based Blockchain. In Proceedings of the International Conference on Vision Towards Emerging Trends in Communication and Networking. IEEE. https://doi.org/10.1109/vitecon.2019.8899393

Paul, T., Buchegger, S., & Strufe, T. (2011). Decentralized Social Networking Services. In L. Salgarelli, G. Bianchi, & N. Blefari-Melazzi (Eds.), Trustworthy Internet, 187–199. Springer. https://doi.org/10.1007/978-88-470-1818-1_14

Prensky, M. (2001). Digital Natives, Digital Immigrants. On the Horizon, 9(5), 1–6.

Raman, A., Joglekar, S., Cristofaro, E. D., Sastry, N., & Tyson, G. (2019). Challenges in the Decentralised Web: The Mastodon Case. In Proceedings of the Internet Measurement Conference (pp. 217–229). ACM. https://doi.org/10.1145/3355369.3355572

Safe Network. (2020). Safe Network: We're Building The New Internet. https://safenetwork.tech

Skaf-Molli, H. (2017). Decentralized Data Management for the Semantic Web. Doctoral thesis. Université de Nantes. https://hal.archives-ouvertes.fr/tel-01618806/document

Solid. (2020). Solid. https://solid.inrupt.com/

Srinivasan, R. (2017). Whose Global Village? Rethinking How Technology Shapes Our World. New York University Press.

Srnicek, N. (2016). Platform Capitalism. Polity, Theory Redux.

Taipale, S. (2019). Intergenerational Connections in Digital Families. Springer.

Tan, J., & Tan, A. E. (2012). Business under threat, technology under attack, ethics under fire: The experience of google in china. Journal of Business Ethics, 110(4), 469–479. https://doi.org/10.1007/s10551-012-1494-0

Tiku, N. (2019). Three Years of Misery Inside Silicon Valley's Happiest Company. Wired, 2019(September), 80–94.

Tor Project. (2020). Browse Privately, Explore Freely. https://www.torproject.org/

Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. Human Resource Development Review, 4(3), 356-367. https://doi.org/10.1177/1534484305278283

Trattner, C., Helic, D., Singer, P., & Strohmaier, M. (2012). Exploring the differences and similarities between hierarchical decentralized search and human navigation in information networks. In Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies (No. 14). ACM. https://doi.org/10.1145/2362456.2362474

van Hardeveld, G. J., Webber, C., & O'Hara, K. (2017). Deviating From the Cybercriminal Script: Exploring Tools of Anonymity (Mis)Used by Carders on Cryptomarkets. American Behavioral Scientist, 61(11), 1244–1266. https://doi.org/10.1177/0002764217734271

Verborgh, R. (2020). Re-decentralizing the Web, for good this time. In O. Seneviratne & J. Hendler (Eds.), Linking the World's Information: Tim Berners-Lee's Invention of the World Wide Web. ACM. (Accepted for publication). https://ruben.verborgh.org/articles/redecentralizing-the-web/

Verborgh, R., Wrigley, S., & Ballardini, R. M. (2019). Decentralizing the Web and the Industrial Internet: an Alternative Solution to Data Control. In R. M. Ballardini, P. Kuoppamäki, & O. Pitkänen (Eds.), Regulating Industrial Internet through IPR, Data Protection and Competition Law (pp. 47-59). Kluwer Law International.

Wild, S., Chudnovskyy, O., Heil, S., & Gaedke, M. (2013). Customized Views on Profiles in WebID-Based Distributed Social Networks. In F. Daniel, P. Dolog, & Q. Li (Eds.), Web Engineering (pp. 498–501). Springer. https://doi.org/10.1007/978-3-642-39200-9_44

Wong, J. C. (2019, March 18). The Cambridge Analytica scandal changed the world – but it didn't change Facebook. The Guardian. https://www.theguardian.com/technology/2019/mar/17/the-cambridge-analytica-scandal-changed-the-world-but-it-didnt-change-facebook

Wood, M. (2019). Today's Firefox Blocks Third-Party Tracking Cookies and Cryptomining by Default. https://blog.mozilla.org/blog/2019/09/03/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/

WWW Foundation (2018). The case for the web. https://webfoundation.org/docs/2018/11/The-Case-For-The-Web-Report.pdf

WWW Foundation. (2020). Contract for the web. https://contractfortheweb.org/

Yeung C-M., A, Liccardi, I., Lu, K., Seneviratne, O., Berners-Lee, T. (2009). Decentralization: The future of online social networking. In W3C Workshop on the Future of Social Networking (pp. 2–7). W3C. http://www.w3.org/2008/09/msnws/papers/decentralization.pdf

Zeronet. (2020). Zeronet: Open, free and uncensorable websites, using Bitcoin cryptography and BitTorrent network. https://zeronet.io/

Zhuo, Z., Zhang, Y., Zhang, Z., Zhang, X., & Zhang, J. (2018). Website Fingerprinting Attack on Anonymity Networks Based on Profile Hidden Markov Model. IEEE Transactions on Information Forensics and Security, 13(5), 1081–1095. https://doi.org/10.1109/tifs.2017.2762825

KEY FACTOR FOR REAL TIME TEACHING
IN HOME OFFICE MODE

Petr Rozehnal, Roman Danel

Department of Computer Science
Faculty of Economics
VŠB – Technical University of Ostrava
petr.rozehnal@vsb.cz, roman.danel@vsb.cz

Keywords

Home office, real time teaching, strategy, process, MS Teams.

Abstract

This paper discusses the impact of a lockdown caused by virus spread on the educational process at a faculty. Using a questionnaire survey among teachers at the faculty, we found out how they coped with this situation, their technical equipment and whether they encountered any problems. The aim of the paper is not to bring an exact evaluation of selected questions, but to show the state of the current situation, to point out possible problems of users and to link these things with the approach to the management of the organization. Based on the identified weaknesses, we bring suggestions and recommendations for improving the process of transition to online learning.

1. Introduction

Groundbreaking events always carry significant changes. It is up to people to take them as a problem or a challenge that moves them in a positive direction. In 2020, we are seeing events of world importance. The spread of coronavirus infection has changed the functioning of the whole society across countries around the world within a few weeks.

In many cases, organizations have switched to home office mode. One of the problems with this change was the speed at which everything happened. The paper deals with some aspects of this transition. The basis is a survey conducted at a selected faculty focused on the humanities education. The Faculty is a part of a public university in the Czech Republic. The research was conducted among academics and focused on the ability to carry out the online learning process without the possibility of being present at the faculty. It is important to mention that this regime was switched over in an extremely short period of time, without the possibility of preparation, and that the online form of education was not implemented at the faculty with regard to accredited forms of education. The pedagogical process has undergone a transformation into a digital form. At the micro level example, the pillars of digital transformation according to Doucek et al. (2017) were documented.

The aim of the paper is to answer predefined questions that are important for the transition and provision of teaching at the faculty. The aim is to analyze the situation and, if necessary, define long-term procedural recommendations that will support work in the home office mode. Although the survey is conducted in an educational institution, the questions have been formulated in such a way that part could be transferred and generalized in any organization. It is clear that most companies and institutions had to deal with a similar situation.

2. Description of the initial situation and definition of topics for assessment

Description of the initial situation and the institution where the survey is conducted. The Faculty carries out education in a form that is typical for the Czech Republic.

The attendance form is dominated by a combined form of study. In the combined form of teaching, full-time teaching is limited, more emphasis is placed on complementing teaching with features typical for distance learning. Full-time education is not replaced in this form by real-time distance online learning. The faculty uses learning management system (LMS) Moodle to support teaching (no matter what form of study), lecture recording system with subsequent export to video and placement in university systems (video, audio and presentation), Office365 for remote work and selected software is available via Internet (remote desktop connection).

The rapid spread of infection in connection with Covid-19 meant the transition to home office within two weeks. Especially for educational processes this meant a significant change in behavior. It can be stated that the university had technologies with which it was possible to implement the teaching process immediately online. Of course, the reality of switching to home office mode was not easy.

The rapid spread of infection in connection with Covid-19 meant the transition to home office within two weeks. Especially for educational processes this meant a significant change in behavior. It can be stated that the university had technologies with which it was possible to implement the teaching process immediately online. Nevertheless, problems had to be overcome because the change brought other demands on equipment and knowledge on the part of workers.

The consequences of this change were mainly reflected in:

- Necessity to have an adequate technical and software equipment.

- The need for a quality Internet connection.

- The need to have the same data available at home as in the university environment.

- Sufficient knowledge to use technologies that support home office work and teaching process applications.

The faculty management strategy was simple. Primarily to replace as much as possible full-time teaching with real time online teaching. Secondary to supplement the study materials with elements typical for distance learning and replace personal form of communication with electronic equivalents.

The primary goal was technically based on Office365 applications, especially Teams (Duffy, 2020), which are available to all University users (academics and students). The second objective was also to be supported by communication in the Teams environment. In addition to traditional teaching materials, offline videos were also deployed, which can be produced individually only with the assistance of a computer, i.e. screen capture software. Based on this baseline situation, we have derived several research goals for evaluation:

- What is the technical equipment in the home office mode? The question is whether academics have the hardware needed for real time online learning?

- What is the availability of software equipment for academics/students?

- What is access to work data in home office? The question is if academics have equal access to working data in home office compared with the normal mode of operation when they are present in the workplace?

- What is the quality of the Internet connection? That means whether the Internet connection does not limit real time teaching.

- To what extent has been implemented the transition to full-fledged on-line education? According to the teaching elements applied in online teaching process.

- Definition of the most common problems, notes and ideas.

3. Data analysis and evaluation

The survey was conducted between the faculty's academic community. Respondents were addressed by targeted email with a link to the questionnaire containing 12 questions. 230 respondents were addressed, the questionnaire returned 26.5 %. In order to assess the return rate, it is necessary to take into account that about 25 % of academics are internal PhD students and colleagues from abroad. These academics are not primarily responsible for course management and their role in teaching is secondary.

The survey was anonymous and was also supported by the faculty management. It was conducted in the fourth week of an emergency situation when academics switched to home office mode. The time was chosen in a targeted way so as to give some time to organize the lessons. The questions were mainly open (or at least with the possibility to write a comment). The reason for this concept of questions was the possibility to express information freely without restricting the preference. In retrospect, we evaluate it positively. Respondents, for example, mentioned software that would be difficult to predict when preparing a questionnaire. Several questions followed each other logically in order to derive the observed aspect. Therefore, only 6 questions are listed and evaluated below.

What is the availability of software equipment for academics/students?

The results of the questionnaire survey indicate that the availability of software for teaching is not a problem at the faculty (note: here is meant the software needed in teaching, not for the organization of teaching). Some comments on the missing software issue were due to a lack of information - the respondents did not know that the software was available under certain conditions (e.g. VPN access).

- 5 % of respondents said they lack certain software (limited-use, licensed software).

- A fundamental problem that was diagnosed on the basis of a questionnaire survey is the answer to the question "Do you know about the possibility of running some programs in the form of remote access?". 29 % of respondents said they did not know about this option and 14 % knew about it but could not use it. Again, this situation is caused by a lack of information about the possibilities and features that the IT infrastructure of the faculty offers.

What is the technical equipment in the home office mode? The question is whether academics have the hardware needed for real time online learning?

As far as hardware is concerned, the situation is more complex than software. Not all respondents have all the necessary hardware.

- 10 % of respondents report insufficient technical equipment (5 % hardware quality and 5 % missing hardware, equipment specially designed for teaching)

- 3 % of respondents report a scanner as missing hardware

- 6 % of respondents mention a missing printer (need to print selected materials with regard to their study, preference to work with the printed version).

What is access to work data in home office? The question is if academics have equal access to job data in home office compared with the normal mode of operation when they are present in the workplace?

Academics are ready for home office mode in the field of storing work data. The investigation did not identify any problems in the availability of data to ensure work continuity. The questionnaire deliberately identified two separate questions about where academics store data and which repositories they access in home office mode.

The basic variants for storing data are as follows: locally on the device, network disk mapped in the faculty environment, cloud storage and portable storage of individual users. Accessing data stored on a local device is trouble-free at home with portable devices and is not possible with classic desktop computers. VPN access is required to access the faculty network drive.

- The results show that only 11 % store data on a desktop computer that they don't have access to from home. At the same time, these users store data on portable devices. Therefore, data availability problems were not identified.

- 55 % of academics have a portable computer on which they store data and have access to it. This hardware is also equipped for audiovisual needs.

- Only three percent of respondents reported a confusing answer when they reported a different location where they store data and other storage to which they have access. However, they have not reported a problem with data access.

What is the quality of the Internet connection? The question is whether an Internet connection does not limit real time learning.

80 % of respondents assess their Internet quality as sufficient and trouble-free. Evaluation is based on subjective perception of connection. For practical reasons, we did not require technical connection specifications. With regard to the fact that workers have been working in home office mode for the third week, we assume that they have a long enough time to assess whether they have noticed a de facto restriction on connection during work.

To what extent has the transition to full-fledged education been implemented online? According to the teaching elements applied in teaching.

The main variants of teaching to assess the degree of transition to real time teaching are:

- 47.5 % of respondents use LMS and traditional materials (presentation, pdf, etc.) only with added elements such as material links, comments, individual tasks.

- 12 % of respondents use LMS, traditional materials and video recordings of lectures - partial replacement of full-time teaching.

- 40.5 % of respondents use LMS, traditional materials and real time video lectures and communication - maximum possible substitution of full-time education with the help of IT tools.

As can be seen, about half of the teachers were able to replace the contact form of teaching.

It is also worth to mentioned:

- 21 % of teachers did not make any changes to the materials / teaching methods, although they did not identify any problems that might be the cause. They responded to the survey (they are not passive), but made no changes in the teaching process (they are passive). We have no explanation for this result.

- 31 % of teachers use different communication tools instead or together with recommended tool by university. 18 percent use Skype, other applications are represented with minimum shares (for example WhatsApp, Facebook, Zoom). In the current situation it is difficult to assess this result. In a positive sense, this is an active approach, probably driven by efforts to

maintain contact with students. In the negative sense, this implies the creation of non-standard communication channels, where it is difficult to maintain a uniform environment and way of communication. If each teacher uses a different means of communication to ensure teaching and communication with students, this situation will complicate the realization of online learning for students.

Definition of the most common problems.

The questionnaire also included an open question where it was possible to state what for academics are missing in the current situation, describe problems and other comments. The most commonly represented suggestions and comments:

- A set of problems caused from the lack of awareness among academics, and the lack of information is due to the worker's fault. These problems would not arise if academics were registering and reading information distributed at the faculty / university. 8 percent of respondents report these issues.

- 15 % of respondents report their requirements for better support e.g. information, instructions, documentation.

- 43 % of respondents did not use the opportunity to write a comment, problem or suggestion.

4. Evaluation of results and recommendations for the future

Retrospectively, it can be assessed as the right decision to leave a number of questions with the possibility of completing in verbal form. Respondents mostly used this option and the results can also reflect information that could not be taken into account in closed questions with limited choice. This also supports the overall survey intention, which was not intended to accurately express percentages, but to provide an overall picture of the situation and how academics dealt with the situation. Surprisingly, there is a relatively low rate of return, as the survey was conducted within one organization and with the support of management.

In the context of the survey's overall focus, about half of the academics switched to active online learning with real-time communication elements, and half remained in the traditional forms (offline materials and communication), which only enriched with other materials, notes, resources, etc. The basic qualitative finding is the fact that switching to forced home office mode can be a trigger event that can positively impact academics' ability to use workflow and groupware applications.

The basic quantitative findings include:

- Total technical equipment is sufficient. The advantage is equipped with portable computers. There are indications that there was a lack of equipment to support a specific pedagogical process (e.g. boards).

- Overall software availability is sufficient. Applications for creating a virtual learning environment are available.

- The problem is the lack of or inadequate form of support e.g. documentation, tutorials, manuals, etc.

- Strong tendency of academics to use other solutions than recommended ones (especially in the area of communication with students).

For organizations, it seems appropriate to develop a strategy that would, in the long term, guarantee the maintenance or strengthening of positives on the one hand and eliminate identified shortcomings

on the other. Given the fundamental role of information technology, much of the recommendations concern IT management.

Specifically:

- Hardware equipment - the use of portable devices has proved to be a positive feature. In the future, the structure of hardware for academics should be considered, with respect to portability.

- Software equipment - it is not only about their availability, but above all about ensuring the ability of academics to use it. See the next point.

- Provide appropriate documentation (beyond the scope of standard documentation). Support must be simple, oriented to the selected issues that are most needed. The speed at which support will be available when needed and its availability is also crucial.

- Implementation of targeted education of workers and their awareness. defining the process of continuous education from familiarization with IS / IT at the beginning of the employee up to updating knowledge based on the development of IS/IT.

- The university should declare the choice of Teams (or generally selected application) as the recommended means of communication and justify this decision (ownership of the Office365 license, stability guarantee, support for group communication ...). When using alternative communicators, 100 % reliability is not guaranteed, see experience using Skype for video conferencing (Danel and Řepka, 2017).

- When deploying preferred solutions and procedures, it is advisable to ensure the use and limit the variability of the resources deployed. A prerequisite is alignment with other parts of the organization and anchoring in the policy of the organization. Managed documentation (for academics and students) should be developed, containing a description of the processes for key areas (e.g. communication, observation, work with the study group, etc.). Currently, a large number of software solutions (often free of charge) are available that tempt you to deploy as "I use what I know" or "I use what I find". This approach raises problems that are well known in the field of IT management.

Despite all the benefits, online learning cannot fully represent the direct contact of students with the teacher. Some subjects require practical demonstrations that are difficult to implement online and interaction with students is also important. Non-verbal communication is completely lacking in on-line learning, for example, when a teacher sees how students respond to a lecture topic if they understand it. Also, the interaction with students on-line is more complicated - some students tend not to ask and attend only formally (Danel et al, 2018). All of this leads to the conclusion that online learning using current IT technologies is an emergency solution and should not be seen as a substitute for classical learning in normal operation after quarantine. However, it is possible to combine standard teaching with digital technologies and to use it for example for lectures by external experts and practitioners.

5. Conclusion

Experiences and knowledge gained during quarantine can be used in the future. The survey focused on the ability to implement the learning process. We start from the fact that the home office regime in the field of day-to-day work is already in operation at universities (and is basically anchored in Czech legislation). The results showed that the knowledge of the key product for online communication and cooperation supported at the university was zero at the beginning of the introduction of quarantine.

It can be used by the academic community in the field of team cooperation on projects, remote consultations and consultancy, to conduct lectures and workshops led by practitioners, external experts or academics from other universities without the need for their physical presence. Possibilities of cooperation on the one hand and the need to be prepared for remote communication and cooperation states (Ministr and Pitner, 2014). Above mentioned need for description of process and documentation is necessary.

Integration into the policy of the organization is a prerequisite for long-term assurance of the benefits of emergency measures and quarantine. Primarily in the areas of human resources management, teaching process and IT management. Once again, the well-known fact that people and their abilities are a key resource has been confirmed. Only if there is an alignment between all the resources necessary to carry out the process can the desired effects be achieved. This confirms the direction in which the latest frameworks work IT frameworks (ISACA, 2019; Open Group, 2020). It is an element of "capability" as an overall expression of the ability to act.

Distribution of information and standardization of procedures also appear to be a major problem. In this paper, we have proposed several measures that could lead to a faster and smooth transition to online learning if a similar situation occurs in the future.

6. References

Danel, R. & Řepka, M. (2017). Experience with teaching using videoconferencing system. Proceedings of IT for Practice 2017, Ostrava, Czech Republic, 191-198.

Danel, R., Řepka, M., Valíček, J., Kušnerová, M. & Harničárová, M. (2018). Education and research in informatics and automation at Faculty of Mining and Geology. Proceedings of CSIT 2018 – Computer Science and Information Technologies, Springer Book, Lvov, Ukraine. Plenary session, XV-XXI.

Doucek, P., Fisher, J., Novotný, O. (2017). Digital economy. Procedings of 25th Interdisciplinary Information and Management Talks – IDIMT 2017, Poděbrady, Czech Republic, Linz: Trauner Verlag Universitat, 33-40.

Duffy, J. (2020). Microsoft Teams Review. [Online]. Available at: https://www.pcmag.com/reviews/microsoft-teams [cited 2020-04-18].

Ministr, J. & Pitner, T. (2014). Towards an ecosystem for academic-industrial cooperation. Proceedings of 22nd Interdisciplinary Information and Management Talks – IDIMT 2014 (Networking Societies - Cooperation and Conflict), Poděbrady, Czech Republic, Linz: Trauner Verlag Universitat, 71-78.

ISACA. (2019). COBIT® 2019 Framework: Introduction and methodology. Schaumburg, USA. ISBN 978-1-60420-763-7.

The Open Group. (2020). TOGAF 9.2 [Online]. Available at: https://pubs.opengroup.org/architecture/togaf9-doc/arch/ [cited 2020-04-12].

GREEN INDUSTRY 4.0 - ANALYSIS OF GREEN ASPECTS PENETRATION IN BUSINESS READINESS MODELS FOR INDUSTRY 4.0

Josef Basl

Faculty of Informatics and Statistics,
PragueUniversity of Economics, Czech Republic;
basl@vse.cz;

Andrea Benesova

Faculty of Electrical Engineering,
Westbohemian University, Czech Republic
benesov2@rek.zcu.cz

Keywords

Industry 4.0, maturity model of industry 4.0, green, sustainability carbon neutrality

Abstract

The article focuses on a closer look at the penetration of green aspects into phenomena called industry 4.0. The analysis is based on the available maturity models for industry 4.0, which are investigated to determine whether and to what extent they fulfill the "green" trends by their name, content or specialized dimensions. This analysis was initiated by a strong interest in integrating the elements and principles of industry 4.0 into enterprises in recent years, on the one hand, and an increased interest in achieving carbon neutral production and society in the coming decades, incl. recent EU Green Deal documents, on the other

1. Introduction

The previous decade, i.e. the period after 2010, was first a period of the end of crisis in manufacturing companies - the mortgage crisis, which subsequently turned into a financial crisis. It was also a period of further deeper focus of companies on the use of information technology. After 2010, there was a wave of increased interest in the connection of IT with the impact on the natural environment, ecology, etc. within their own production, use in companies, but also the final stages of their liquidation. This trend was referred to as "green IT" and was thus focused on two application concepts, both "green by IT" and "green of IT" (A Green Knowledge Society, 2010, Greener and Smarter ICTs,2010).

Only a few years later, after 2013, the development and deployment of IT in companies began to be declined in another key "revolutionary" context. The term "industry 4.0" or broader as the 4th Industrial Revolution has become a key term for IT penetration into enterprises (Deloitte. Industry 4.0, 2015).

Both these important trends emphasizing the application of IT in companies took place essentially independently of each other, but with a partial overlap. From today's perspective, however, both these trends should be bee understood together with mutual interaction and support.

Therefore, if we depict them together in a two-dimensional space, where the "X" axis would be the trend of achieving a higher degree of maturity in the direction of "industry 4.0" and the "Y" axis would then capture the trend of increasing "green IT" parameters, clearly three routes depicted in Fig. 1:

- path 1 is heading towards improvement in the direction of "industry 4.0",

- path 2 in the direction of 'green IT',

- path 3 thus means that changes towards industry 4.0 with the positive green impacts.
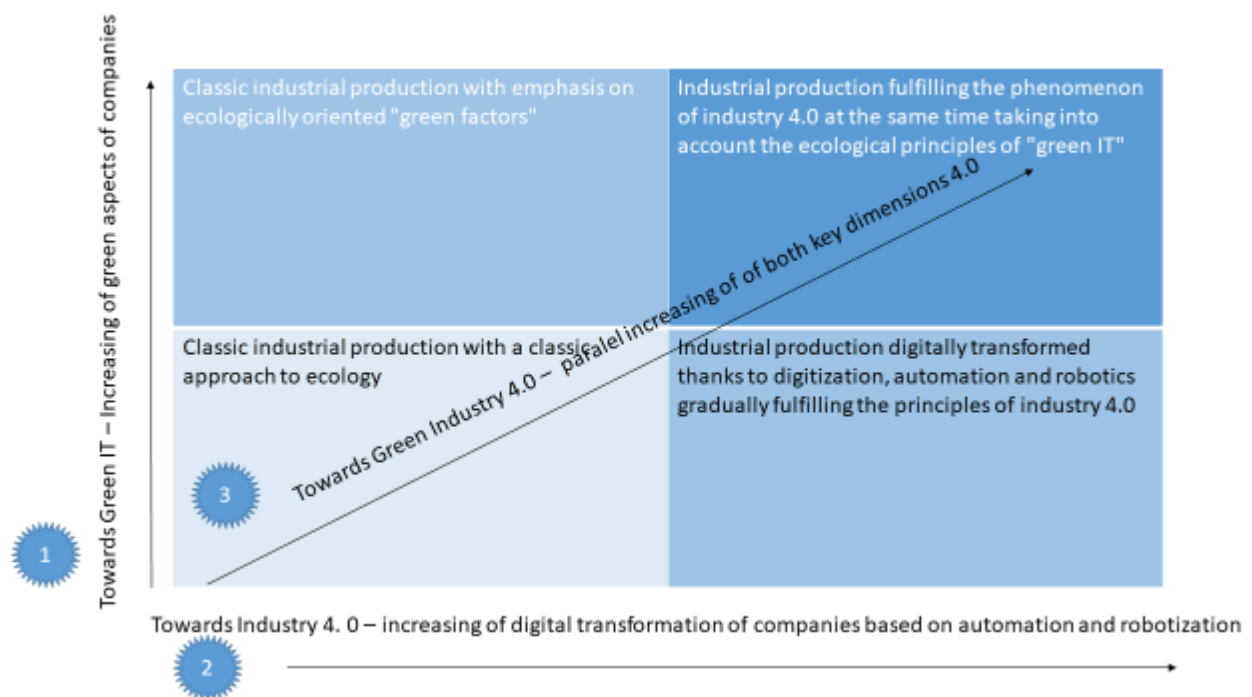
Fig. 1 Transformation of classic industrial production towards industry 4.0 and green and possible synergy (combination) in "green industry 4.0"

From the point of view of the authors of this article, it is currently important that both trends take place simultaneously and will be implemented together. It is this intersection that is referred to in this article as "Green Industry 4.0, and this article is dedicated to.

2. Green Manufacturing 4.0 – current evolution phase in industry

The term "green Industry 4.0" is not yet precisely defined. However, its concept can be approached through a separate view of the "industry 4.0", followed by a separate view of "green", and this can be inferred by their intersection with "green industry 4.0".

2.1. Reflection of IT trends in industry 4.0

Recently in the last decade, the IT trends in industry has become more and more popular and it is known "Industry 4.0". This trend (sometimes called the fourth industrial revolution) certainly affects

or will affect most people and most companies. Impacts of such previous revolutions have manifested themselves in various areas - not only in industry itself, but also in science, health, ensurance, agriculture and trade for example. It is likely to be assumed that Industry 4.0 will have similar effects to the previous industrial revolutions (Fassman and Šulc, 2016).

Many significant consultancy companies such as Gartner Group (Gartner, 2016), BCG or Deloitte (Deloitte, 2016) present their typologies of what the current trend towards Industry 4.0 should include. It is clear that the concept of Industry 4.0 is based on industrial integration mediated by information technology. This integration involves real-time or near-real-time data sharing, information sharing, and continuous communication.

In general, the concept of Industry 4.0 can be characterised as a transformation of production as separate automated factories into fully automated and optimised manufacturing environments. Production processes are linked vertically and horizontally within enterprise systems. To that end, sensors, machines and IT systems are interconnected within the value chain across the enterprise boundaries. For this purpose, the Cyber-Physical System (CPS) is the cornerstone for smart factories. These factories have the ability to autonomously exchange information using the Internet-based communication protocols, thus responding in real time to potential mistakes and adapting to changing customer demand for products. Smart products create clever products that are uniquely identifiable.

2.2. Role of IT in sustainable trends towards green deal

IT sector has also had a significant influence on the effectiveness and competiveness of enterprises and generally on the whole society for last decade. IT sector is generally understood like a "key enabler" of the economic growth but it is also consider as an important factor of sustainable "green growth". Nowadays, the innovation of IT is more and more used in this way. "Green" behaviour of the IT itself is not less important.

But IT is not influencing the enterprises, society and households. The IT sector to be able to offer its products and services itself needs relatively huge amount of energy. IT sector depends heavily on energy, the most of which comes from non-renewable resources such as fossil fuel. IT can improve environmental sustainability and contribute to greener trends in a number of ways. Heavily used IT equipment includes data centres, servers, PCs and printers. Many companies have therefore dramatically reduced power consumption by adopting efficient ICT technologies, efficient processors and virtualization of equipment.

There are some examples to illustrate this situation. The energy consumption of IT is really enormous. It is expected that IT industry for example accounts for approximately 2% of global carbon dioxide (CO_2) emissions (Gartner 2007; Marwah et al. 2009b). About 75% of the ecological footprint comes from the use of ICT products and 25% embodied carbon in IT products. The embodied carbon is emitted during extraction, manufacture, distribution and operation of IT products. The IT carbon emission is estimated to grow at 6% each year until 2020 (The Climate Group 2008).

This current growth in energy consumption and carbon emission puts IT industry in an unsustainable condition. The ICT industry needs to proactively consider life cycle analysis of its products and innovate to reduce environmental impact. ICT organizations have started to take environmental sustainability as a part of corporate social responsibility.

In the IT sector, sustainability could be achieved by better management of the IT products and accessories to use less energy, and hence emit less carbon. ICT companies can manage their retired computers and other electronic products by handing over to others for reuse instead of holding them in storage. They should find environment-friendly way to recycle PCs, cell phones and other electronics.

The sustainable paradigm shift is based not only on new technologies like IT. The changes are not imaginable without the changes of behaviour of producers and consumers of IT. The IT is really bringing about a fundamental shift in economic and social lives. It has already changed the way how producers and consumers interactively communicate.

2.3. Green Industry 4.0 – useful intersection of digital transformation and green deal

Green Industry 4.0 represents an important intersection, but also the simultaneous application of both trends. The changes towards "industry 4.0" should thus also be changes positively supporting "green" impacts. The introduction of elements and principles of "industry 4.0" should deteriorate the impact on the environment and the negative impact on climate change in no way.

The increased interest in "green" issues is confirmed by the EU strategic document "Green Deal" (The European Green Deal, 2020). However, we should not forget that under the presidency of industrialized Sweden in 2010, the green perspective was integrated into the European strategy then called the "Green Knowledge Society"(A Green Knowledge Society, 2010) .

3. Analysis of green attributes in the current industry 4.0 readiness maturity models

3.1. Methodology of analysis

The declaration of the main trends of focusing on "industry 4.0" and "green" is just the beginning. The starting point for this can be the industry 4.0 maturity models that would follow and support such green dimensions in the construction of industry 4.0. The analysis should provide the necessary information on this. This article uses existing models and analyses to find answers to the following research questions:

1) Is green principle sufficiently reflected in existing literature?

2) Do existing maturation models capture aspects of green and, if so, adequately?

3) If they capture these models, is it in the form of a separate dimension or just a sub-dimension usually assessed in such models?

The analysis is done by searching and comparing in existing models whether they contain:

- Information about green in the model name.

- Information about green not in the title but in the overall description and approach of the model.

- Green information within a separate green dimension name.

- Information about green not in the title, but in the content of a non-green dimension.

3.2. Data applied for analysis

The analysis is based on data representing available and sufficiently methodically described maturity models. Currently, there are more than 100 maturity models that are designed not only for the overall preparedness of companies for Industry 4.0, but also for individual technologies. Most maturation models contain dimensions and grades of so-called levels. However, some models contain not only dimensions but also sub-dimensions that specify the dimension in more detail. Maturity models differ mainly in context (Industry 4.0, Cybersecurity etc.), number of dimensions, sub-dimensions and

number of levels. Since there are a large number of maturity models, only those that are most relevant in terms of evaluation should be selected for more detailed analysis.

For the analysis 27 maturation models were chosen based on the journal where the article was published, the impact factor of journal, the number of citations and the year of publication. The models were selected in the time period 2011-2020. The characteristics of the analysis data include - quantity, focus, scope, detail, year of publication, authors (university, consulting company, national methodologies, etc.). The 27 maturity models analysed, are described in Table 1.

4. Results of analysis of green aspects in industry 4.0 readiness models

Most of the models are focused on Readiness of company to Industry 4.0, Digital Business Maturity or Cybersecurity Assessment. The analysed data show that most maturation models contain 5 dimensions and 5 rating levels. The technology dimension contains mostly questions focused on currently used technologies and further which other technologies the company plans to implement. In dimension strategy and leadership most of the models evaluate to what extent is Industry 4.0 established and implemented in company's strategy, but do not assess the sustainability of the company or the impact on the environment.

Table 1. - Results of analysis of maturity models in terms of Green Aspects

Name	Year/Author	Green in Name	Green in Context	Green in Dimen-sions	Green in Sub-dimensions
Evaluation form for evaluation of digital maturity of the company	2015/ Zdeněk Havelka et al.	x	x	y (Environment and Digital Management, Technology)	x
Manufacturing Readiness Index	2011/Jung et al	x	x	x	x
RAMI 4.0 (Reference architecture model for Industry 4.0)	2015/ BITKOM, ZVEI a VDMA	x	x	x	x
IMPULS	2015/Lichtblau et al.	x	x	x	x
Guidelines for I 4.0	2015/Anderl et al. – VDMA	x	x	x	x
The Industrie 4.0 Component Model	2015/ZVEI	x	x	x	x
Digital maturity of the enterprise	2016/MPO	x	x	x	x
SIMMI 4.0 (System Integration Maturity Model Industry 4.0)	2016/C. Leyh	x	x	x	x
The Connected Enterprise Maturity Model	2016/Rockwell Automation	x	x	x	x
Industry 4.0 / Digital Operations Self-Assessment	2016/Geissbauer et al. - PwC	x	x	x	x
Industry 4.0 Maturity Model	2016/Schumacher et al.	x	x	x	x
Maturity Model (Considering Diversification as Enabler)	2016/Ganzarain & Errasti	x	x	x	x
A Categorical Framework of Manufacturing for Industry 4.0	2016/Qin et al.	x	x	x	x

Study Based Analysis on the Current Digitalization Degree in the Manufacturing Industry in Germany	2016/Bogner et al.	x	x	x	x
Development of an Assessment Model for Industry 4.0: Industry 4.0-MM	2017/E. Gökalp et al.	x	x	x	x
Maturity Model for Data Driven Manufacturing (M2DDM)	2017/Weber et al.	x	x	x	x
Digital Readiness Assessment Maturity Approach (DREAMY)	2017/Carolis al.	x	x	x	x
A Smartness Assessment Framework for Smart Factories Using Analytic Network Process	2017/Lee et al.	x	x	x	x
Industrie 4.0 Maturity Index (ACATECH)	2017/Schuh et al.	x	x	x	x
Maturity and Readiness Model for I 4.0 Strategy	2018/K.Y.Akdil et al.	x	x	x	x
Towards a Framework for Assessing the Maturity of Manufacturing Companies in Indsutry 4.0 Adoption	2018/Scremin et al.	x	x	x	x
Maturity model for manufacturing enterprises	2018/Andreas Schumacher	x	x	x	x
Contextualizing the outcome of a maturity assessment for Industry 4.0 – 360 Digital Maturity Assessment	2018/M.Colli et al.	x	x	x	x
Concept for an evolutionary maturity based Industrie 4.0 migration model	2018/Stefan et al.	x	x	x	x
Logistics 4.0 Maturity in Service Industry	2019/KarolinaWerner-Lewandowsk	x	x	x	x
The degree of readiness for the implementation of Industry 4.0	2019/Pacchini et al.	x	x	x	x
Cybersecurity in the context of industry 4.0	2020/A.Corallo et al.	x	x	x	x

The key result is that with the exception of one model, all other analysed maturity models do not include a dimension that assesses the environment or other green aspects.

However, green aspects will be an integral part for companies in future with a view to the new Digital Strategy and Green Agreement of European Union. For this reason, a new green dimension should be added to a maturity model: either as a cross-sectional dimension or as one of the sub-dimension of other existing dimensions

5. Conclusion:

Manufacturing companies, as mentioned in the introduction to this article, are influenced in their management not only by technical progress, but also by the market situation and the situation in the society. Their decisions are thus influenced by such events like financial crises, the risks of terrorist attacks, migration waves, and climate change and, more recently, the current impact of the global pandemic coronavirus. These incentives can have different effects on innovations in the digital transformation of companies, and these innovations can be associated with one of the phenomena (green or industry 4.0), which they then focus on.

However, this may not be the only approach because, as innovations are usually not deployed separately (for example, a combination of two or more phenomena such as IoT, edge computing, block chain, business analytics or artificial intelligence). Thus "industry 4.0" with "green" seems to be a very useful combination. It is important that "industry 4.0" from the start will support the green view, it could be only neutral but in no case, the building of "industry 4.0" should have environmental negative effects. This article should have pointed out this fact. The following research surveys will focus on the impact of "4.0" elements on the "green balance".

6. Reference:

Deloitte. Industry 4.0: Challenges and solutions for the digital transformation and use of exponential technologies [online]. 2015 [cit. 2017-03-13]. Available: https://www2.deloitte.com/content/dam/Deloitte/ch/ Documents/manufacturing/ch-en-manufacturing-industry-4-0-24102014.pdf

European Commission: Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions- A resource-efficient Europe – Flagship initiative under the Europe 2020 Strategy, Brussels, COM(2011) 21 (2011)

European Commission. Coordination of European, national & regional initiatives. In: European Commission [online]. 2016 [cit. 2017-03-13]. Available z: https://ec.europa.eu/digital-single-market/en/cordination-european-national-regional-initiatives

EU Report: A Green Knowledge Society An ICT policy agenda to 2015 for Europe's future knowledge society, A study for the Ministry of Enterprise, Energy and Communications, Government Offices of Sweden by SCF Associates Ltd Final Report (2010)

DJSI 2007: Dow Jones Sustainability Indexes (DJSI), Dow Jones Sustainability World Indexes Guide. http://www.sustainability- ndex.com/djsi_pdf/publications/Guidebooks/DJSI_World_Guidebook_90.pdf (2007)

Gartner - Top 10 Strategic Technology Trends for 2016, http://www.gartner.com/technology/research/top-10-technology-trends/, Accesed: 2016-04-27

Global Trends 2030: Alternative Worlds, National Intelligence Council, 2012, https://globaltrends 2030.files.wordpress.com/2012/11/global-trends-2030-november2012.pdf Accessed: 2016-04-27

Greener and Smarter ICTs, the Environment and Climate Change, OECD Report (2010)

Mahaffy. M.: Defining Sustainability, Washington State University, http://www.arch.wsu.edu/09%20publications/ sustain/defnsust.htm (1999)

National Economic Committee: The national framework of competitiveness strategy, Czech government, Prague (2011)

Premier of the State Council of China and K. Q. Li, "Report on the work of the government," in Proceedings of the 3rd Session of the 12th National People's Congress, March 2015, Accessed: 2016-04-27.

Report: Digital Transformation - Alma Mater 4.0, http://www.cebit.de/en/news/latest-news/alma-mater-4.0.xhtml, Accesed: 2016-04-27

Recommendations for implementing the strategic initiative INDUSTRIE 4.0, 2013, http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonde rseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf. , Accessed: 2016-04-27

Rupert J. Baumgartner, J. R., Ebner, D.: Corporate Sustainability Strategies: Sustainability Profiles and Maturity Levels Sustainable Development Sust. Dev. 18, 76–89, Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/sd.447 (2010)

The European Green Deal, https://www.gndforeurope.com/, 2020

Tondl. L.: Science, technology and society. The modern tendency and transformation of interrelations, Czech Academy of Science, Prague, (in Czech), 184 pp. ISBN 80-7007-051-X (1994).

United Nations Educational Scientific and Cultural Organization (UNESCO), Indicators of sustainability reliable tools of decision making. Available at: http://unesdoc.unesco.org/images/0015/001500/150005e.pdf. (2006)

WCED, World Commission on Environment and Development (WCED), Our common future. Oxford: Oxford University Press. http://www.un-documents.net/wced-ocf.htm (1987)

THE ROLE OF THE YOUNG GENERATION IN FAMILY BUSINESSES

Klára Antlová, Petra Rydvalová, Marián Lamr

Technical University of Liberec
Klara.antlova@tul.cz, petra.rydvalova@tul.cz, marian.lamr@tul.cz

Keywords:

Innovation, family business, succession, young entrepreneurs, interest of young students the entrepreneurship, economic sectors, industrial district

Abstract

This article examines the relationship of the young generation to the family business in the context of the innovative performance of the parents' businesses where they work. Another essential factor is the sector in which businesses operate. This factor is important for the topic of research on the impact of the family business on the traditional industrial districts (ID) sector. The industrial district is understood as a network of small and medium-sized enterprises with some specific knowledge, education, tradition, and skills. Agglomeration of SMEs is very often highly specialized. The results of the present survey point to the importance of sharing knowledge across generations.

1. Introduction

Family businesses are very often based on continuity across the generations and these young family members could be a source of innovation. They are resolute and ready to face the challenges ahead, can help to maintain the family character of the business and keep the family values alive. A new generation of young, highly motivated, and mostly well-educated men and women are ready to take over the leadership of their family businesses. They can help to grow the company in a rapidly-changing economic and business environment, and they also could be essential in commercial challenges as digitization. In line with the development of industry and global society, digitization is becoming part of our lives. Its importance has now been practically demonstrated by the COVID virus crisis (Sahni, 2020). In line with the current situation, there is a clear need to change thinking about the organization of the economy towards long-term sustainability for future generations. There is a possibility of "return" to the importance of the local economy, an integral part of which the family businesses are. These often continue in the traditional industry and thus take advantage of the existence of so-called industrial districts.

Modern technology, E-commerce and use of information and communication technology (ICT) are growing and enable new possibilities of staying in touch with customers and these young members in families are a key part of this innovative and education processes. This could be seen also in the history of some oldest businesses. Ernst and Young and the University of St. Gallen developed the Global Family Business Index which provides significant insights into the world's largest family-owned businesses, ranked by revenues. The largest family-owned businesses in the world include many global companies (such as Fiat, Nestle etc.). Unsurprisingly, in this age of transformation the Family Business Index this year sees a significant increase in family businesses that are leveraging new technologies. The survey from Deloitte (2016): "Next-generation family businesses - Evolution

keeping family values alive" proved that the main challenge is getting the previous generation to acknowledge that the next generation can take over running the family business. Figure 1 explains how the next generation of the family business is becoming the leaders of the biggest challenges that will be maintaining family values, succession planning and introducing further professionalism and innovations into the family firms. This article aims to find the importance of sharing knowledge across the generations in the family business.

Data from the long term research (GUESSS, 2018) indicate that just small part of the students is ready to work in the family businesses after graduation (see Antlova, Rydvalova and Lamr, 2017). But the familiniess of the firm and tradition are important factors in the development of so-called industrial districts and clusters. The authors, therefore, investigate (due to the possible impact of industrial districts on the competitiveness of companies) if there are some differences in the economic sectors.

The organisation of the article is following, next part of the article brings literature research, third part explains how the research data were obtained, in this part is also the validation of the research question and in the last part of the article, the results are discussed.
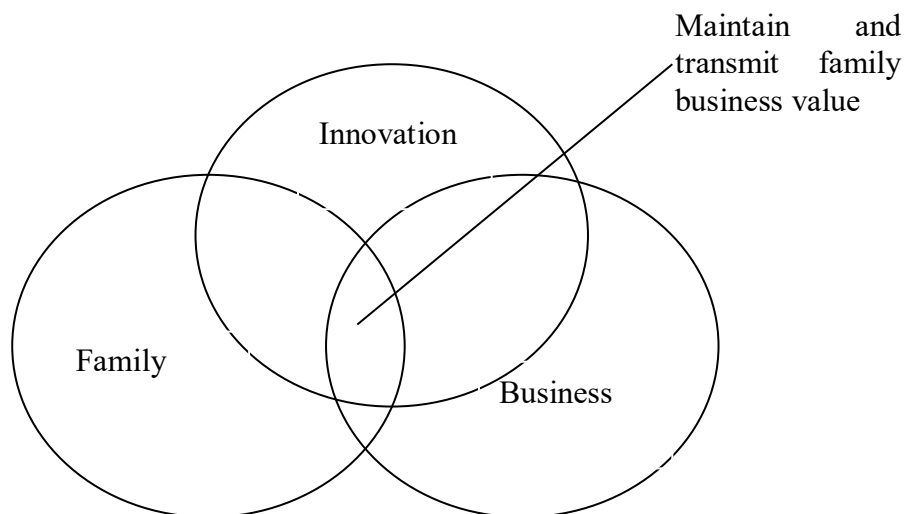
Figure 1 Relationship of the young generation to the family business (Deloitte Next-generation family businesses, March 2016)

The authors of this article have long experience in the family entrepreneurship, their problems and also with the nascent young generations. In the north of the Czech Republic, the area fulfils the definition of the industrial district (according to Marshall), the glass making industry has a long tradition; it goes back more than seven hundred years. This area assembles many glassworks, studios, museums and schools in the Bohemian Paradise, Lusatian Mountains, Jizera Mountains and the Giant Mountains. This creates a new opportunity not just for producing but also for tourism. In connection with this opportunity, for example, the Regional Authority of the Liberec Region prepared a project for the support of the glass industry called "Crystal Valley". Nowadays a lot of these companies are using modern communication channels, using modern internet marketing and still offer very traditional processing of glass and glass components. Today we can see that this tradition of old glass processing is still very attractive for young entrepreneurs. One of the entrepreneurs who is working in parents business said: "I could smell the glass factory from childhood". Also, Marshall (2013) explains that "something is in the air" (1890, reprinted 2013). He calls this a hereditary skill. It means

that the young generation in the family is naturally learning. He later changes the word "something" for the word "atmosphere". Also, Becattini in 1990 described this family atmosphere as creative. Marshall (2013) also explains that when some economic sectors are based in some special area (a good example could be the glass sector in the Liberec region "Crystal Valley") and, likely, it stays here a for a long time. In this context, we can talk about the so-called industrial district. This relates to the development of education and specialization in the region.

2. Literature review

Tradition plays an important role in family businesses, but that alone is not enough. Without innovation, companies will not be competitive. This is a big challenge for young members of family businesses. The authors De Massis et al. (2013) in the article: "Innovation through Tradition: Lessons From Innovative Family Businesses And Directions For Future Research" called this strategy as innovation through the tradition.

Some other articles compare family and non-family businesses as Craig and Dibrell (2006), De Massis et al. (2015), Gudmundson et al. (2003), as well as Pittino and Visintin (2009), Craig and Dibrell (2006) and they mention that family businesses are more open to using modern channels for communication with young members, they have much more informal decision making and are flexible in their processes and it all leads to a more innovation-friendly atmosphere.

Marshall also explains why the companies have profit from the industrial area, it means that the benefit from the concentration of producers and economic atmosphere in a particular area is bringing more customers (1890, reprinted 2013, pgs. 225 – 227). This advantage of industrial districts or zones is also mentioned by F. Belussi and K. Caldari (2009). The use of highly specialised machinery and hereditary skills could be a very important factor in the local market. This is the case for instance of the above mentioned Glass Valley in Northern Bohemia. This Industrial district (ID) is understood as a network of small and medium enterprises in some special knowledge, education, tradition and skills. Very often the agglomeration of SME is highly specialised. We should not forget that also the social conditions are very important for these districts, the innovations here are more incremental than radical changing trends. However, today we can see that the use of ICT is rising in many areas because of the Covid crisis. On the other hand, still, sharing of the generations' knowledge is a necessary and very important factor. Especially in the family enterprises, the young members are educated in some family surrounding and this can help to raise the interest in entrepreneurship and also in some families traditional specialization. The support of parents, social changes and new technological innovations (Industry 4.0), these all are the necessary conditions for future growth and competitiveness - Magretta (2011, s. 159). According to Cucculelli, Storai (2015) the industrial district is the most important factor for the growth in the existing family business. Their research proved that trust between the companies (called reputation capital) reduces transaction costs and also implicates the positive externality.

There are three important dimensions of family influence that should be considered: power, experience, and culture. These three dimensions or subscales are called F-PEC model, which creates an index of family influence. This index enables comparisons across businesses concerning levels of family involvement and its effects on performance as well as other business behaviour. Authors of this index are Klein, Astrachan and Smyrnios (2005). Rutherford and et al. (2008) studied 831 family businesses and concluded that F-PRC index explains more the scale of family influence than the growth of the company. This influence is very important for strategic decisions (Habbershon et al., 2003) or (Rutherford, M.W., Kuratko, D.F., & Holt, D.L. (2008). The extent of the family's influence on the company's strategic decisions cannot avoid even in the innovation process. Calabro et al. (2018) conducted an extensive systematic review of innovation in family firms, by analyzing more

than 100 peer-reviewed articles in journals published between 1961 and 2017. In their research, they identified the main gaps in this topic. One of them was the question of family history, time dimension and tradition, how the use of past knowledge affects the introduction of new products and the return on innovation. Miller et al. (2014) deal with the topic of the time dimension of the formation of organizational behaviour in family companies. Based on knowledge from family dynamics, identity theory and social identity theory, they point out how early family experiences can affect engagement, motivation, but also conflicts in family businesses. This, in turn, affects the transfer of the company to later generations. This research also deals with the issue of perception of the family business by the potential young generation, in terms of innovation and industry.

3. Data and methodology

The author's research team is a part of a big international survey (since 2016) established in 2003 at the University of St.Gallen (Switzerland). Every 2–3 years, a global data collection effort takes place. The eighth wave was in autumn 2018, with 52 countries participating in the survey. The text of this article presents selected data for all countries from the GUESSS survey in 2018. This led to a dataset with 208,636 responses from students from more than 1,000 universities (GUESSS, 2018). For every participating country, there is one responsible country team, which coordinates data collection in that country. The goal of this research is to discover unique and new perspectives on student entrepreneurship in the university environment. The research explores several topics, such as Entrepreneurial intentions, Nascent entrepreneurship, Growth and performance of new ventures and Family firm succession. The questionnaire has 15 pages and the authors in this article focused on the role of young members in family businesses, which are considered an important factor in the competitiveness of industrial districts and clusters. Therefore, the data are further used within the project GA18-01144S, see acknowledgement.

The following research question identifies the involving interest of young generation in the development of family businesses and later also concerning the type of industry and business sector.

- Is there dependency between young members of the family businesses and the innovativeness of the company?

First, all respondents who did not answer any of the monitored questions were removed. After this check, 50,055 records of respondents remained for evaluation. It means the students who answered that their parents have their businesses. Also the amount of 21, 495 (43%) students answered that they are working in the parent's companies. An important factor was therefore whether the respondents considered their parents' company as a family one. For the explorations of this research question, the authors used one of the survey questions:

- "Do you regard your parents' business as a family business?"

From the 50,055 records altogether 27, 582 (55%) respondents answered yes. Another important parameter was whether respondents could have up-to-date information about the company. So the question was whether the students work in the family company. 16,159 (59%) students said that they are working with their parents in the family business. We use the Chi-square test of the independence to determine the relationship between these categorical characters (see Table 1).

A null hypothesis is that there is no correlation between the family businesses where young family members are working and the rate of the performance of the companies. The alternative hypothesis is that family business is innovative when the young members are working there. The observed P-value, which defines the minimum level of the significance for rejecting the null hypothesis, is 0.000 (the min. level of significance is 5%), so the null hypothesis is rejected (Table 1). There is a correlation between the fact that students working in family businesses consider their parents'

businesses more innovative in comparison with their competitors. We also used the Cramer coefficient of contingency. Here, the value of the Cramer coefficient is 0.105, which means that the dependence exists but it is weak.

Table 1 Symmetric Measures

Symmetric Measures		Value	Asymptotic Standard Errora	Approximate Tb	Approximate Significance
Nominal by Nominal	Chi	0.105			0.000
	Cramer's V	0.105			0,000
Interval by Interval	Pearson's R	0.102	0.006	17.048	.000c
Ordinal by Ordinal	Spearman Correlation	0.103	0.006	17.114	.000c
N of Valid Cases		27582			
a. Not assuming the null hypothesis.					
b. Using the asymptomatic standard error assuming the null hypothesis.					
c. Based on normal approximation.					

The authors Dess & Robertson (1984), Powell, Eddleston et al. (2008) explains that family business owners can create organizations that uniquely meet their needs, allowing them to pursue personally defined goals and strategies. Within the GUESSS questionnaire in 2018, the following indicators were monitored in this context: 1 Sales growth; 2 Growth of market share; 3 Profit growth; 4 Job creation; 5 Innovation. Another question in the survey which we used for the verification of our research question was:

- "How do you rate the innovativeness performance of the business compared to its competitors over the last three years?"

In this case, the Likert scale was used ("1" means much worse and "7" means much better). The results are in the next Figure 2.
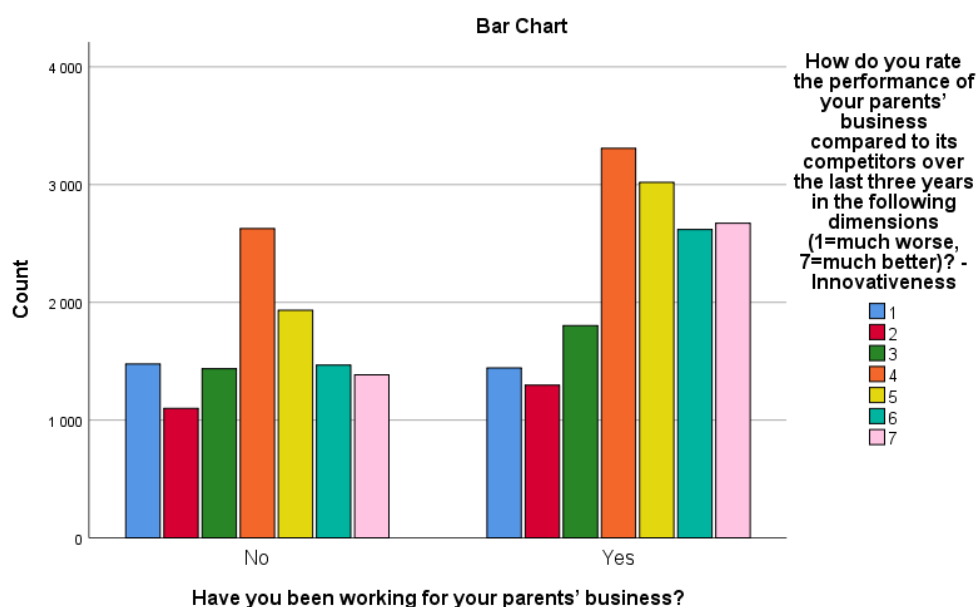
Figure 2 Perception of the performance of family businesses

The authors also searched the innovativeness according to the sectors. This is asked in the survey question:

- "Which sector is your parents' business mainly in?"

Next Figure 3 demonstrates the results, so we could see that especially in the trade business the companies are the most innovative. This could be connected with using information and communication technology for communication with customers or suppliers using the opportunity of e-commerce.
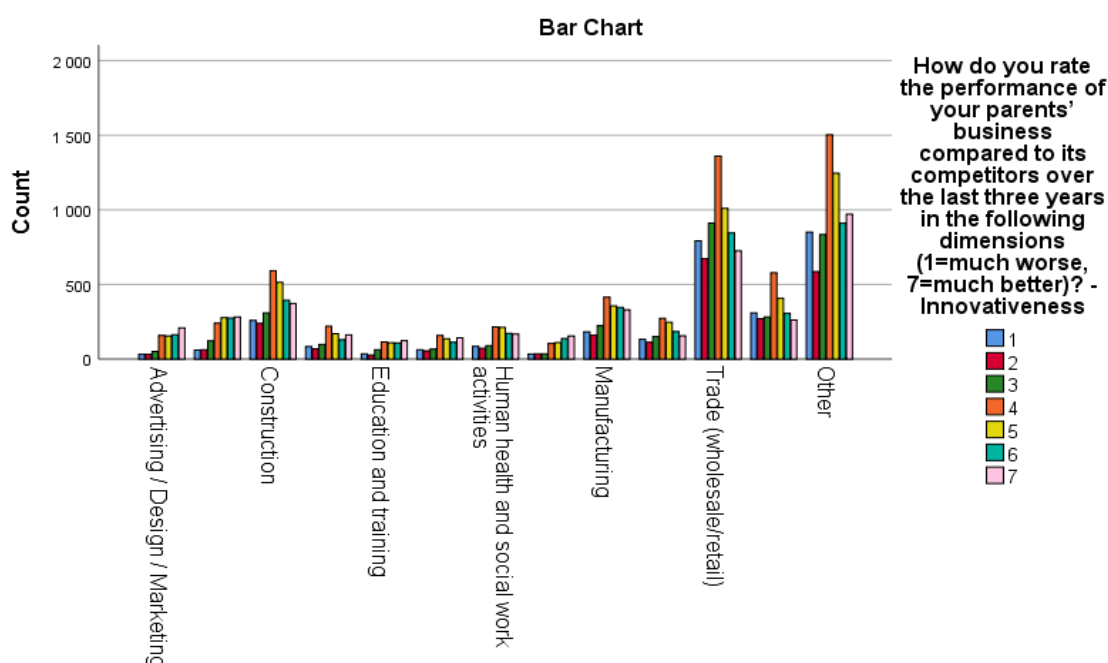
Figure 3 Results of the differences in sectors

4. Discussion and conclusion

Results of this research show the importance of the young members in family businesses and this demonstrates that they are one of the key factors for the innovation processes. Therefore the Czech universities started new courses focused on family business management. This new generation of young, highly-motivated and mostly well-educated men and women is ready to take over leadership of their family businesses. They are resolute and ready to face the challenges ahead. This is consistent with the research of Miller et al. (2014).

In connection with this, we can conclude that the involvement of the young generation is very important for innovativeness in family businesses. This is better in those sectors that have more scope (or are rather easier) to involve young people in implementing innovations in the digital field, such as trade. The path of engaging the young generation in IS / IT is the right one and can be very important for maintaining the family business, traditional industries and, as a result, the development of industrial districts in the region. This topic will be discussed in further author's research.

5. Acknowledgement

6. Reference

Antlová, K., Rydvalová, P. a Lamr, M. 2017. Vztah studentů k podnikání 1. vyd. Liberec: Technická univerzita v Liberci, 2017 ISBN: 9788074943799.

Becattini G. 1990. Alfred Marshall e la Vecchia Scuola Economica di Cambridge, Il Pensiero Economico: temi, problemi e scuole, 1990. Turin UTET (pg. 275-310).

Calabrò, A. et al. 2018. Innovation in Family Firms: A Systematic Literature Review and Guidance for Future Research. International Journal of Management Reviews, 21(3), 317–355.

Chrisman J. et al. 2008. Toward a Theoretical Basis for Understanding the Dynamics of Strategic Performance in Family Firms. Entrepreneurship Theory and Practice.

Cucculelli, M. Storai, D. 2015. Family firms and industrial districts:: Evidence from the Italian manufacturing industry, Journal of Family Business Strategy, Volume 6, Issue 4, December 2015, Pages 234-246.

Craig, J. B. L., & Dibrell, C. 2006. The natural environment, innovation, and firm performance: A comparative study., Family Business Review, 19, 275-288.

De Massis, A., Frattini, F., Pizzurno, E., & Cassia, L. 2013. Product innovation in family vs. non-family, firms: An exploratory analysis. Journal of Small Business, Management. Doi.org/10.1111/jsbm.12068.

Deloitte. 2016. Next-generation family businesses [online]. Available from: https://www2.deloitte.com/.

Dess, G. G., & Robinson, R. B. 1984. Industry effects and strategic management research. Journal of Management, 16(1), 7-27.

Franz W. Kellermanns, Kimberly A. Eddleston, Ravi Sarathy and Fran Murphy. 2012. Innovativeness in family firms: a family influence perspective. In Small Business Economics. Vol. 38, No. 1, The Entrepreneuring Family (January 2012), pp. 85-101 (17 pages). Published by: Springer. Available from: https://www.jstor.org/stable/41410003.

GUESSS. 2018. Global University Entrepreneurial Spirit Students' Survey. Switzerland. Available from: http://www.guesssurvey.org/.

Habbershon, T., Williams, M., & McMillan, I. 2003. A unified systems perspective of family firm performance. Journal of Busi-ness Venturing, 18(4), 451-465.

Gudmundson, D., Tower, C. B., & Hartman, E. A. 2003. Innovation in small businesses: Culture and ownership, structure do matter. Journal of Developmental Entrepreneurship,8(1), 1-18.

Kuratko, D. F., & Hodgetts, R. M. 2001. Entrepreneurship, a contemporary approach (5th ed.). Florida: Hartcourt College Publishers.

Magretta, Joan. 2011. Understanding Michael Porter: The Essential Guide to Competition and Strategy. 2011-12-20. Boston: Harvard Business Review Press. 256 pgs.

Marshall, A. 2013. Principles of Economics. London: Macmillan. Peter Groenewegen, reprinted (1890/2013).

Miller, D. & I. Le Breton-Mille. 2014. Temporal considerations in the study of family firms. Eu-ropean Journal of Work and Organizational Psychology,23, pp. 669–673.

Pittino, D., & Visintin, F. 2009. Innovation and strategic, types of family SMEs. A test and extension of Miles and Snow's configurational model. Journal of Enterprising Culture, 17, 257-295.

Powell, G. N. and Eddleston, K. A. 2008. The paradox of the contented female business owner. Journal of Vocational Behavior, Vol. 73 No. 1, pp. 24-36.

Sahni, D., Takacs, B., Pannala, K. Bertran, J. 2020. The impact of COVID-19 on SMB IT demand, [online], Available:https://marketing.analysysmason.com/acton/attachment/3183/f-32561605-b09a-4b12-8f3d-d03841cfb942/Analysys_Mason_COVID_SMB_IT_May2020.pdf

PREDICTING CUSTOMER LIFETIME VALUE IN PERFORMING ARTS EXPLOITING NON-TRANSACTIONAL DATA

Martin Potančok, Lubomír Štěpánek, Jiří Mařík, Lucie Šperková, Marie Vítová Dušková

Faculty of Business Administration
University of Economics, Prague
martin.potancok@vse.cz, lubomir.stepanek@vse.cz, jiri.marik@vse.cz,
lucie.sperkova@vse.cz, marie.duskova@vse.cz

Keywords

Customer classification, Customer lifetime value, CLV, Machine learning, Neural networks, Random Forests, Recency, Frequency, Monetary, RFM

Abstract

Nowadays, customer lifetime value (CLV) is a key concept in direct response marketing, supporting innovation and data-driven decision-making in many business areas. It can be described as an approach to assessing the financial value of each customer of a given customer base. Most of the currently used techniques employ transactional data and assume a contractual setting. The aim of this article is to propose a way of predicting CLV in a narrow field of performing arts – concert organizing. We employ an under-used design, creating a model of discrete transactions utilizing non-transactional features, such as categorical data of the performed events. The non-transactional data draw on the assumption based on long-term observation of customers' behavior that specific properties of the concerts might influence the decision whether to attend a concert in the future. Exploiting non-transactional data makes this work different from other CLV-focused papers. Our approach utilizes recency, frequency and monetary (RFM), and combines this data with other relevant non-transactional attributes about the concert audience behavior. We compare the predictions of traditional machine learning algorithms: multinomial logistic regression, naive Bayes classifiers, decision trees, random forests, support vector machines as well as neural networks. The results show that the role of non-transactional data is not entirely unambiguous: most models worked better without the non-transactional data. Thus, the non-transactional data contribution to the predictions of CLV is not entirely straightforward.

1. Introduction

Predicting customer lifetime value (CLV) is traditionally considered a regression task: an income from a given customer in a specified future period is estimated based on various customer data (e.g. historical transactions, customer rankings, etc.). Although this concept is heavily employed in internet marketing, it is highly neglected in public and artistic fields as well as in other specific environments. This article aims to propose an approach to predicting CLV in a narrow field of the performing arts – concert organizing.

The proposed method is based on real-life data from a small non-profit organization that organizes chamber music concerts in two concert series. Their concert dramaturgy is often non-traditional, featuring unusual combinations of musical instruments, the fusion of different musical genres, and the combination of music with dance and spoken word. At the same time, they try to create concert opportunities for young beginning artists who can perform alongside renowned top performers. Although the concerts take place in a historic hall in the city center frequently visited by tourists, the audience is domestic. In the past ten years, the organization has gained steady and loyal supporters, which allowed us long-term monitoring of the audience's behavior.

First, we mention several characteristics of the performing arts, describing the problem of CLV classification in this environment, and define requirements for the proposed approach. Second, we discuss current methods related to our investigation. Afterwards, the data used for the modeling is described and the proposed models are implemented. Finally, we discuss the results and derive managerial conclusions.

1.1. **Problem Definition**

Classical music management is characterized by a generally conservative approach to marketing activities with only a few exceptions in the marketing mix, especially regarding communication. Although cultural institutions (organizers of classical music concerts) try to communicate with their audience using modern technologies and sometimes novel means, they often do so without more in-depth analysis of their target groups and customers. Our research shows that these organizations typically use only basic means to gain knowledge about their audiences, most often questionnaire surveys (their own primary data collection), if at all. Unlike in other fields (e.g. e-shops, banking, telecommunications), analyzing customer data is very sparse due to the absence of evidence-based solutions and the very low awareness of management of the benefits of such essential means as CLV. According to this description, we come up with the following assumptions for our approach:

- Concepts with simple interpretation and simple applicability are preferred.
- Cultural institutions deal with lower amounts of data (in comparison to e-commerce).
- Attributes used for prediction should be easy to obtain, broad, and understandable to classical music management.

From a CLV perspective, this environment is a non-contractual, always-a-share setting with repeated transactions on a discrete-time basis. Transactions can only occur at fixed regular intervals associated with specific events; however, in our study, spending varied throughout the period as ticket prices to the event increased. According to Pfeifer and Ovchinnikov (2011), if firms can only serve a limited number of customers, the decision to serve a customer runs the risk of displacing another customer in the future.

Concerning the above, we consider CLV a classification task rather than a regression task, since it provides results that cultural institutions can use in decision making. Seeing CLV as a regression task enforces subsequent grouping of customers who have a similar numeric CLV value: from a marketing communication point of view, organizations want to communicate with segments of customers, rather than individuals.

In this paper, we focus on the classification of customers (audience of performing arts institutions) according to their CLV into three different classes (low, middle, high), which correspond to the estimated income from a customer in the upcoming year. The classification is based on features arising from customer transaction data (customer ID, event date, transaction value) and derived from them, and also attributes of the concerts, i.e., non-transactional data such as genre, style, popularity of a performer etc.

2. Related work

CLV has been widely studied for more than a decade. In this section, three approaches related to our investigation are described: (1) RFM methods, (2) machine learning algorithms, and (3) probabilistic models.

Recency, frequency, monetary value (RFM) models are simple scoring models well described in (Fader, et al., 2005a), which are used in direct marketing for targeting customer segments to increase the response rate. *Recency* indicates time elapsed since the last purchase; *frequency* stands for a total number of purchases; and *monetary* is the value of the total generated income per customer. Customers are classified into segments based on the values in every of the three categories. Using this principle, RFM is implicitly linked to CLV, as it enables to predict future behavior. However, Jasek and Vrana (2014) demonstrate that recency and frequency are key to predicting CLV, but only for one period, and that monetary value may not determine future transactions. Different refined versions of RFM appear in literature, e.g. LRFM (L standing for the length of the relationship) (Alvandi et al., 2012) and RFMC (C standing for the clumpiness of the customer purchases) (Zhang et al., 2015).

Other important approaches to CLV are based on data mining, machine learning and nonparametric statistics, the predictive ability being utilized by using five different algorithms: (i) projection-pursuit models; (ii) neural network models; (iii) decision tree models; (iv) spline-based models (generalized additive models – GAM), multivariate adaptive regression splines – MARS), classification and regression trees – CART); and (v) support vector machines (SVM). Most of these can analyze a much larger number of variables, mainly when variability is too large to estimate the parameters. Gupta (2006) suggested how to improve prediction performance by combining various models: (a) machine learning with bagging; (b) econometrics with forecasts; and (c) statistics with weighting the predictions of different models.

Probabilistic models, also known as buy-till-you-die (BTYD) models, use statistical stochastic processes to predict whether an individual will still be an active customer in the future. The most widely used model, often serving as a benchmark in non-contractual settings, is Pareto/NBD (Negative Binomial Distribution) invented by Schmittlein, et al. (1987). The model describes repeat-buying behavior in a non-contractual context. Pareto/NBD was further discussed and modified by other researchers (e.g. Fader, et al., 2005b; Platzer et al., 2016; Jasek, 2015, Jasek, et al., 2019). Implementations of the probabilistic models can be found in the BTYD R package (Dziurzynski, et al., 2014) and BTYDplus (Platzer, 2016).

3. Methods

3.1. Data Description

Our model is based on repeated transactions occurring in discrete time, which means that the transactions are associated with specific dated events. The processed data comes from a small non-profit organization putting on chamber music concerts in Prague, Czech Republic. The data spans the period of 2012–2017. Based on the data from 2012–2016, predictions regarding 2017 were made.

There are two distinct groups of features used for predictions (which are also physically stored in two different datasets): data about the transactions (and data derived from them), and non-transactional data about the concerts connected to the transactions.

Each row of the transactional data contains one customer ID, a date of the attended concert, the number of purchased tickets, and the price per ticket. The transactional dataset includes 2,268 rows belonging to 693 unique IDs (customers), each row containing one transaction. The following features were derived from this data:

- Recency (R), frequency (F), and monetary (M) values for each customer and an RFM score calculated as a weighted average using weights of equal sizes (⅓ for R, F, M) – obtained from the transaction data from 2012–2017.

- An annual amount, which is an average value of money spent yearly by each customer in the period 2012–2017 – obtained by dividing their total expenditures in 2012–2017 by the number of years when at least one purchase occurred. While the annual amount uses the number of years with at least one transaction (i.e., maximum six years), the monetary is calculated over the entire period of 2012–2017 (i.e., six years).

To ensure reliability of the calculations, the number of expected transactions as well as the recency and frequency values for each customer were computed using the BTYD package in the R language and environment (Dziurzynski, 2014).

The non-transactional data contains categorical data on performer(s) popularity (well known, moderately known, unknown), major and minor music styles (baroque, classicism, romantic era, late romantic era, jazz, cross-over, 20th century), concert type (recital, chamber, music and dance, music and spoken word) and the day of the week when the event took place (Monday to Sunday). This specific non-transactional data was selected based on long-term observations of customer behavior and the resulting assumption that only the above attributes possibly influence audience's decisions whether to attend an event in the near future or not. The non-transactional data includes 51 different events taking place in the period of 2012–2017; each event's record consists of the values of the above attributes.

In order to deal with the categorical data consistently, we added a number of ratio-based features into the dataset. For each value v of a categorical attribute a (of an event) for a given customer c and a given period p we computed the following:

- A proportion of the number of concerts with an attribute a of a value v attended by the customer c in the period p to the total number of concerts attended by the customer c in the period p:
  - E.g. if a customer attended the total of 12 concerts in the period of 2012–2016 and 3 of them were of the baroque major era concert style, then their proportion of the baroque major era concert style equals 0.25.
  - The variable's label was created using the following form: ratio_of_<value v>_to_all_<attribute a>_levels_within_ID.

- A proportion of the number of concerts with an attribute a of a value v attended by the customer c in the period p to the number of all concerts organized in the period p
  - E.g. if the total of 12 concerts of the baroque major era concert style were held in the period of 2012–2016 and a customer attended 3 of them, then their proportion of the baroque major era concert style equals 0.25.
  - The variable's label was created using the following form: ratio_to_all_<value v>_of_<attribute a>_within_dataset.

While the number of features derived from the transactional data is low (just four: recency, frequency, monetary, and their average value, RFM), the number of features obtained from the non-transactional data is relatively high (equal to the summation of all the levels of all the above attributes, i.e., 18,

multiplied by two for each of the proportions above). The output class was defined based on the exploration of the transactional data, namely the annual amount (as defined above) of money spent at the festival (a sequence of all the concerts). The borders of the classes were determined using terciles of the annual amount in the following way: low = the annual amount is lower than the first tercile (about 50 CZK); middle = the interval between the first and the second tercile (about 50–150 CZK); and high = customers who spent more than the value of the second tercile (about 150 CZK) in the period of 2012–2016.

Eventually, the final dataset was split into a training set and a testing set. While the training set contained all the data from the years of 2012–2016, the testing set only covered 2017. As mentioned above, the models aimed at predicting the classification of each customer based on their estimated CLV in 2017. In other words, they learned to assign each customer into one of the CLV classes (low, middle, or high). The performance of the models was measured using confusion matrices and prediction (classification) accuracies.

We assumed that there are two distinct customer subsets. First, customers who attend a high proportion of all the organized concerts and due to their high loyalty are less sensitive to changes in their attributes. And second, customers who only attended one or a few concerts in the past and are not much interested in ongoing events. Therefore we tried to define a subpopulation of customers who attended an average proportion of all the organized concerts, assuming that those customers are mostly attracted by one or more particular attributes of the concerts; in such a case, the addition of the non-transactional data could improve the prediction accuracy. The accuracies were calculated using two training sets: (i) the training set including all the data for 2017; and (ii) the training set including only the data for the customers who attended between 25% and 75% of all the concerts in 2017.

3.2. Models Implementation

Since the volume of the (training) data was relatively low, we employed traditional machine-learning approaches like support vector machines, Naive Bayes, etc., rather than deep learning approaches, focusing on the following list of methods: (1) multinomial logistic regression; (2) naive Bayes classifiers; (3) support vector machines; (4) random forests; and (5) backpropagation neural networks. All the calculations were performed in the R language and environment, using the *e1071* and *rpart packages*, and the *randomForest libraries*.

Multinomial logistic regression draws on multivariate linear regression, but uses a link function, usually the logit function, to predict posterior probability that an observation belongs to a class {low, middle, high}. In reality, two sub-models of binary logistic regression comparing a focal class to a reference one are learned and the final class is determined by maximizing the posterior probability (Wright, 1995).

Naive Bayes methods are machine learning algorithms based on applying Bayesian techniques (Zhank, 2004). Bayes' theorem is described by the following relationship where y in {low, middle, high} is dependent variable and x1 through xn are feature vectors:

$$P(y \mid x_1, \ldots, x_n) = \frac{P(y)P(x_1, \ldots, x_n \mid y)}{P(x_1, \ldots, x_n)}$$

An observation is classified into a class (low, middle, or high) with the maximum posterior probability calculated using the formula above.

Support vector machines (SVM) is another simple machine learning algorithm that can be used for both regression and classification. However, it is applied more frequently to classification problems. The principle behind SVM is to apply a kernel transformation on raw data projected onto an appropriate hyperplane and to find a plane optimally discriminating the data between the two

classes. Thus, three sub-models have to be learned, followed by a voting scheme classifying into three classes {low, middle, high}. This algorithm is preferred due to its accuracy and lower computing demands (Vapnik, 2013).

Random forests are typical classification algorithms. They work with a collection of tree-structured classifiers with randomly bootstrapped rules in their decision nodes, $\{h(x, k), k = 1, ...\}$ where the $\{k\}$ are independent and identically distributed random vectors, and each tree casts a unit vote for the most popular class at input x. (Breiman, 2001).

Neural networks were developed to simulate the network of neurons in the human brain. The algorithm computes a function of the inputs by propagating the computed values from the input neurons to the output neuron(s) and using the weights as intermediate parameters that are adjusted iteratively in a backpropagation fashion. Learning occurs by changing the weights connecting the neurons. (Aggarwal, 2018).

We repeated the models' learning and subsequent classifications two times: (i) first considering all predictors, i.e., both the transactional (recency, frequency, monetary, RFM score) and the non-transactional (the derived ratio-based attributes) variables, and (ii) second, considering only the transactional variables.

Besides the random forest model building, an importance analysis using a similar algorithm was performed to measure the relative importance of each of the predictors in the model by calculating a mean decrease of Gini's index and of accuracy. While the mean decrease of Gini's index calculates each variable's importance as the sum over the number of all tree splits that include the variable proportionally to the number of samples it splits, the mean decrease of accuracy computes the variable's importance as a difference between the accuracy of a random forest model including the variable and a random forest model replacing the variable with random noise (Hastie, 2001).

Assuming that outputs from probabilistic (BTYD) models, closely related to the predicted CLV classes, could cause information leakage in the process of classification, we decided not to include them in the machine-learning models as predictors. Besides the classification accuracy used to measure prediction performance, we used 10-fold cross validation to increase the robustness of the accuracy estimates.

4. Results and discussion

Table 1 shows the prediction accuracies for the test data including all the observations for 2017. Table 2 shows the prediction accuracies for the dataset consisting only of customers who attended between 25% and 75% of the concerts in 2017. For each of the models, the accuracy is reported as a mean of ten partial accuracies coming from ten iterations of the 10-fold cross-validation. The training set for the 10-fold cross validation was invariant and covered the period of 2012–2016; the test set was subset randomly ten times over the year 2017. The tables show results for both the above designs, including (left columns) and not including (right columns) the non-transactional data.

Considering the results in Table 1 (all the data for 2017 in the test set), the random forest model learned without the non-transactional data returned the best predictive accuracy (about 0.990), while the most inferior accuracy was achieved by a naive Bayes classifier using the non-transactional data (about 0.326). Surprisingly, all the models achieved better prediction accuracy when they did not take into account the non-transactional data. This could be explained by the fact that the transactional data is much closely linked to the output variable, i.e., the classes of CLV in 2017; thus, when combined with the predictors, the non-transactional variables could act as noisy variables not

bringing any useful information into the prediction, or they could even make the predictions worse in comparison to when only transactional data was used as predictors.

Table 1. Accuracies in 10-fold cross validation considering all the data for 2017

	With non-transactional data	Without non-transactional data
Multinomial logit regression	0.664	0.826
Naive Bayes	0.326	0.822
Support vector machine	0.437	0.763
Random forests	0.838	0.990
Neural networks	0.373	0.457

Source: (authors)

Considering the results in Table 2 (the test set including only data about customers attending between 25% and 75% of the concerts in 2017), the prediction accuracies are generally higher than in the previous design. This corresponds to our assumptions and could mean that the non-transactional predictors could improve the accuracy when we are interested in the customers attending an average amount of concerts (i.e., not all, but not so few). The highest accuracies were also achieved by random forest models (about 0.936 with the non-transactional data and 0.990 without the non-transactional data). It is worth noting that the classification accuracy of neural networks is higher when the network does work with the non-transactional data (0.626) than when it does not (0.607). This corresponds to our expectations that considering the non-transactional data should increase its predictive accuracy.

Table 2. Accuracies in 10-fold cross validation including only data about customers attending

between 25% and 75% of the concerts in 2017

	With non-transactional data	Without non-transactional data
Multinomial logit regression	0.534	0.712
Naive Bayes	0.602	0.716
Support vector machine	0.610	0.916
Random forests	0.935	0.990
Neural networks	0.626	0.607

Source: (authors)

The reason why some of the models perform better when employing the non-transactional data than others could be researched more theoretically. Both neural networks and random forests can differentiate between the effects of individual predictors. That may explain why the prediction accuracy is similar or even better when the non-transactional data was taken into account compared to when it was not.

Neural networks use their learned weights to indicate which of the predictors are more significant for the output: the higher the weights of the connections between the individual inputs and neurons of the first hidden layer, the more significant are the inputs for the model. By applying this principle to the design using the non-transactional data as inputs, the learned weights of the connections

mentioned above are probably low as they do not affect the output predictions much. Thus, the prediction accuracy could be higher (or similar) for the model using both the transactional and the non-transactional data than for the model employing only the transactional data, as we can see in the last row of Table 2 (and Table 1).

Similarly, random forests can also differentiate between how individual predictors may influence the output prediction; this depends on a bootstrapped-based random selection of predictors in each of the trees of the forest and each of the nodes of a tree. Therefore, the significance of the individual predictors for the final prediction accuracy may vary a lot, and considering the non-transactional data might not affect the accuracy much (Table 1 and 2).

On the other hand, algorithms such as multinomial logistic regression, naive Bayes classifiers, and support vector machines are sensitive to each predictor at the input. While multinomial logistic regression may be biased by noisy predictors (in this case the non-transactional data) correlated to other ones due to multicollinearity, naive Bayes classifiers assume that all predictors are independent from one another. None of these approaches could be adequately applied to the real (transactional and non-transactional) data. Analogously, the support vector machine in this basic use also has to generate the prediction using all predictors since support vector items are enumerated for all the predictors. Furthermore, the support vector machine as a native dichotomous classifier demanding linear separability of the observations' universe usually fails when applied to a complex high-dimensional problem with highly correlated predictors. More noisy variables (that cannot be avoided by algorithm crunching), i.e., the non-transactional data in this case, tend to lower the prediction accuracy (Table 1 and 2).

Figure 1 and Figure 2 show the importance of the variables based on the mean decrease of accuracy and of Gini's index, respectively. As we can see, the monetary and the frequency (i.e., the predictors based on the transactional variables) are the most important variables in the models, followed by the derived ratio-based variables (i.e., the predictors based on the non-transactional variables). The numerical difference between the joint impact of the monetary and the frequency and all the other variables is significant and could imply the transactional variables rather than the non-transactional ones are far more important for model accuracy.
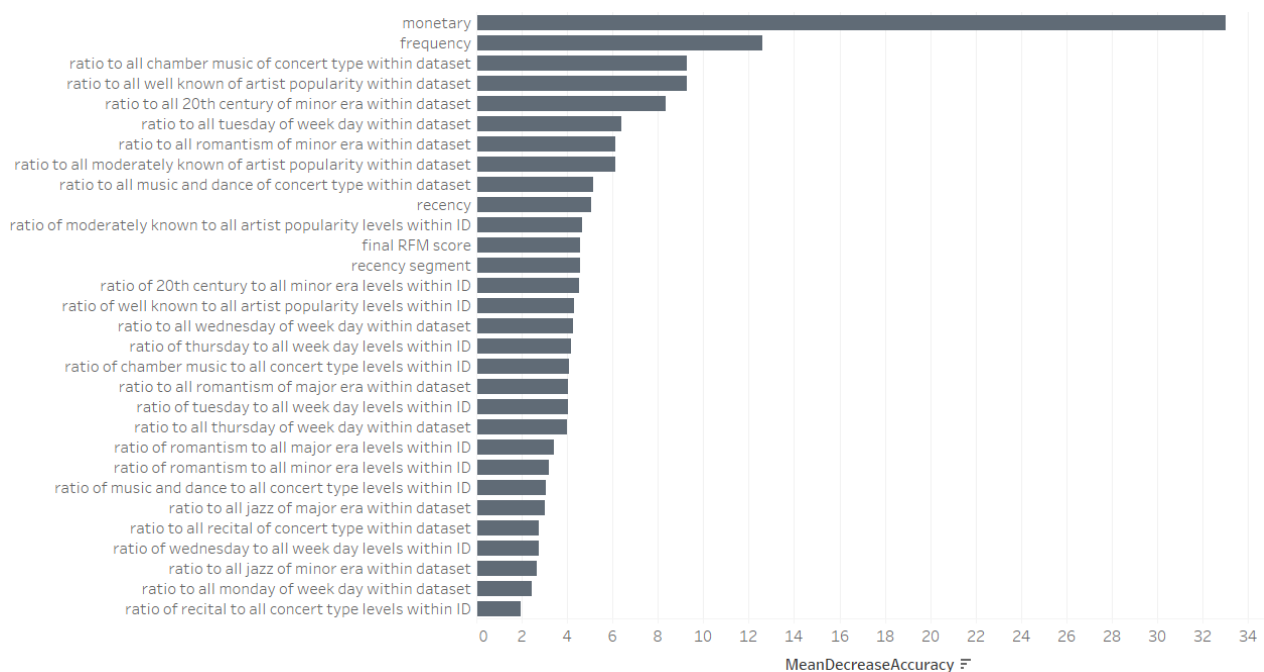
Figure 1. A mean decrease of accuracy (authors)
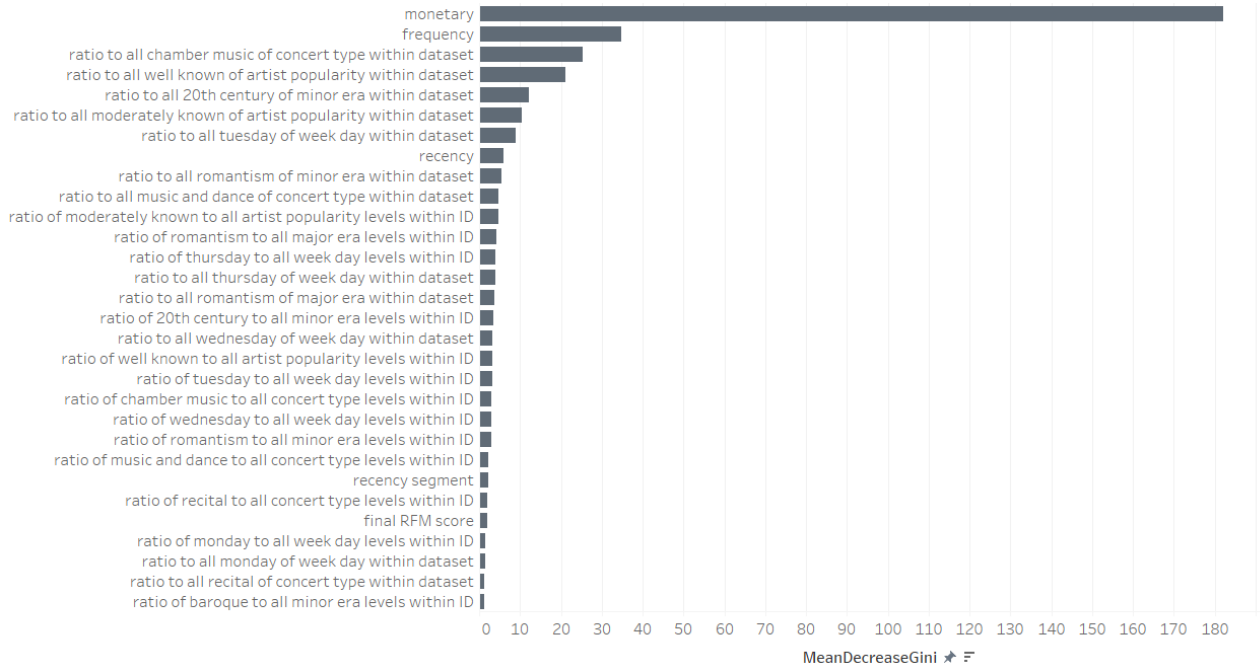
Figure 2. A mean decrease of Gini's index (authors)

5. Conclusion

The results of the customer lifetime value computations which in addition to the traditionally employed transactional data also use non-transactional attributes (such as music style of the concert and others) could be relevant for specific combinations of predictive models and subpopulations of customers, particularly neural networks applied to "median" attendees (those who attend an average number of events rather than the most loyal and the most sporadic customers). The outputs can help the organizer to predict the attendance of various concerts and adjust the dramaturgy based on this prediction; although the models used in our research worked better without the addition of the transactional data, their performance when including this data (dramaturgy, music style, invited artists) was still sufficient to be useful for the management of the organization. Consequently, they could better plan concert programs and at the same time better target marketing communication with its customers. They could, for instance, create special offers for specific customer segments and adjust their activities in audience building.

Such data-supported strategy can affect future sales and hence the economic stability of the organization. Besides, as the organization only uses its existing data, there are no additional costs related to acquiring other customer data or individual primary data collection.

This paper proposes a novel method for estimating CLV classes based on a combination of transactional and non-transactional features. We have shown that under certain conditions the non-transactional data can help achieve better results when combined with the traditional transactional ones than when a model only uses transactional data, especially when focusing on average customers attending roughly a half of all events. The refining of CLV in categorical fashion, i.e., establishing disjunct CLV classes instead of modelling such a variable continuously, followed by employing classification tasks rather than regression ones is innovative in itself. Considering CLV in this fashion could be easier and more useful when making managerial decisions, as communication with segments of customers (based on the disjunct classes) is more natural than with each customer individually.

Further work aims at gathering more training data to investigate deep learning approaches. We will also analyze the results using metrics other than accuracy.

6. Acknowledgement

7. References

Aggarwal, C. C. (2018). Neural networks and deep learning. Springer, 10, 978-3.

Alvandi, M., Fazli, S., & Abdoli, F. S. (2012). K-Mean clustering method for analysis customer lifetime value with LRFM relationship model in banking services. International Research Journal of Applied and Basic Sciences, 3(11), 2294-2302.

Breiman, L. (2001). Random Forests. Machine Learning 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Dziurzynski, L., Wadsworth E., & McCarthy D. (2014). BTYD: Implementing Buy 'Til You Die Models. R package version 2.4. url: https://CRAN.R-project.org/package=BTYD.

Fader, P. S., Hardie, B. G., & Lee, K. L. (2005a), Rfm and clv: Using iso-value curves for customer base analysis. Journal of marketing research 42(4), 415–430

Fader, P. S., Hardie, B. G., & Lee, K. L. (2005b). Counting your customers the easy way: An alternative to the Pareto/NBD model. Marketing science, 24(2), 275-284.

Gupta, R., & Sichei, M. M. (2006). A BVAR model for the South African economy. South African Journal of Economics, 74(3), 391-409.

Hastie T., Thibsirani R., & Friedman J. H. (2001). The Elements of Statistical Learning. 2nd edition. Springer Verlag.

Jasek, P. (2015). Advances in performance management using customer equity. In: IDIMT 2015: Information Technology and Society - Interaction and Interdependence - 23rd Interdisciplinary Information Management Talks, 349–356.

Jasek, P., Vrana, L., Sperkova, L., Smutny, Z., & Kobulsky, M. (2019). Comparative analysis of selected probabilistic customer lifetime value models in online shopping. Journal of Business Economics and Management, 20(3), 398-423.

Jasek, P., & Vrana, L. (2014). Managerial impacts of different computation models for customer lifetime value for an e-commerce company. 8th International Days of Statistics and Economics, 678–686

Pfeifer, P. E., & Ovchinnikov, A. (2011). A note on willingness to spend and customer lifetime value for firms with limited capacity. Journal of Interactive Marketing, 25(3), 178-189.

Platzer, M., & Reutterer, T. (2016). Ticking away the moments: Timing regularity helps to better predict customer activity. Marketing Science, 35(5), 779-799.

Platzer, M. (2016). BTYDplus: Probabilistic Models for Assessing and Predicting your Customer Base. R package version 1.0.1. url: https://CRAN.Rproject.org/package=BTYDplus.

Schmittlein, D. C., Morrison, D. G., & Colombo, R.: Counting your customers: Who-are they and what will they do next? Management science 33(1), 1–24 (1987)

Vapnik, V. (2013). The nature of statistical learning theory. Springer science & business media.

Wright, R. E. (1995). Logistic regression. In L. G. Grimm & P. R. Yarnold (Eds.), Reading and understanding multivariate statistics (p. 217–244). American Psychological Association.

Zhang, Y., Bradlow, E. T., & Small, D. S. (2015). Predicting customer value using clumpiness: From RFM to RFMC. Marketing science, 34(2), 195-208.

TOWARDS 2030 – DIGITAL TRANSFORMATION IN CRISIS MANAGEMENT

INTEROPERABILITY AND CRISIS MANAGEMENT IN PANDEMIC OUTBREAK SCENARIOS

An Overview on different Case Studies of the Austrian Approach to tackle the SARS-CoV-2 spread and adherent management challenges.

Karin Rainer, Alois Leidwein, Peter Nemenz, Markus Hoffmann

Austrian Agency for Health and Food Safety (AGES)
karin.rainer@ages.at

Georg Neubauer

AIT Austrian Institute of Technology GmbH
georg.neubauer@ait.ac.at

Georg Aumayr

Johanniter Research
Georg.aumayr@johanniter.at

Abstract

The global spread of the novel Corona Virus, SARS-CoV-2, in the first half of 2020, resulted in a massive strain of multiple and diverse crisis management structures all over the world. Interoperability of governmental and non-governmental responder organizations as well as of analysis and communication infrastructure were put under a real stress test and concepts of the response to this pandemic outbreak were rapidly set up, modified, changed, and applied according to multiple integrated and changing data sources. Austria, as a country with a relatively mild development of the COVID effects, will be selected as a use case provider to exemplify lessons learned. This paper will give a comprehensive overview of the dynamic development of the initial and follow up steps. Examples from a public service provider AGES, the Austrian Agency for Health and Food Safety, the research organization and technology provider AIT, and the first responder and social services organization Johanniter Unfallhilfe, will show lessons learned, good practice, but also challenges and open questions or gaps in the wake of this complex crisis scenario.

1. Introduction and global background

The spread of the novel Corona Virus, SARS-CoV-2, causing the different forms and levels of the disease commonly referred to as COVID-19, can be globally declared as the major critical event and coordination challenge of the first half of the year 2020.

At the beginning of the global spread it was rather unclear, what the origin of this novel virus was exactly. Besides this question that is still not answered without scientific dispute, it was also for a longer time not totally clear how infectious this new and rapidly transferred virus was that was spread by global travelers around the world. Also, the mortality rate was heavily discussed until up to now, an open question due to the rapid mutation of the virus. Therapy and vaccination were of course not available for this emerging and broadly broadcasted threat (Ludwig 2020). In addition to these open and pressing issues, it took a seemingly long time to get valid and broadly accepted information about the dynamically changing situation. Even the World Health Organization (WHO) was hesitant in the first phase of the development – and seemingly for an undue period after (see Mackenzie 2020) – to officially confirm a pandemic outbreak. However, COVID-19 was declared a global pandemic by the end of March 2020 despite the first, also politically and economically triggered protraction.

The global will, top down measures and mostly also bottom up compliance (s. SailLabs 2020) to minimize the spread were initially not very successful. The virus was developing and spreading in record time, measures of locking down, hygiene measures and the "social distancing" – soon to become a key word of the Corona-crisis – were for some countries either coming too late, not taken severely enough, or by other means not sufficient to stop the local health care services from being over-burdened by intensive care patients with COVID-19. The circle of new infections and of the transmission of the virus was running on a very high level as statistics show (WHO 2020 and Roser 2020). The environmental vicious cycle (s. Biswaranjan 2020 and Figure 1) of asymptomatic carriers, international travel, a lack of basic hygiene and measures on the one hand and severely affected COVID-19 patients on the other hand was a driver for the development.
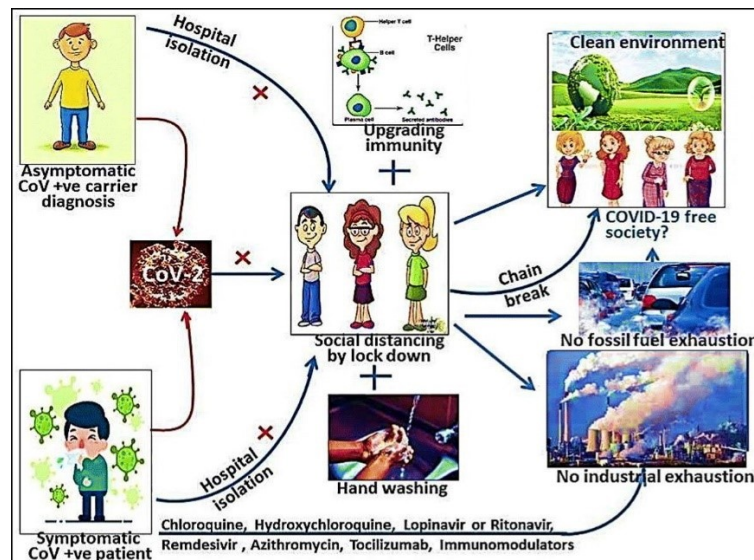
Figure 1: Graphical Overview of the CoV-19 environment and perspectives (Biswaranjan 2020)

The development in Austria was after the severe outbreak in the neighboring regions in Italy initially supposed to be moving in parallel with two weeks delay (see e.g. Complexity Science Hub Vienna 2020).

This paper will shed light from several perspectives on the challenges, gaps and open questions but also the requirements identified in generic crisis management context. Basing on these examples and

the findings gathered in former research and practical appliance of tackling critical developments it will be shown, that can be adapted and thus applied for managing novel and dynamic crises as the COVID-19 event.

2. Coordination and Communication Challenges – use case of the public service provider AGES in Austria

AGES, the Agency for Health and Food Safety in Austria, a public service provider under the ownership and supporting the Federal Ministries in charge of Health and Agriculture, was a major player and responsible communication platform in the so-called Corona-Crisis.

From the very beginning of the global spread of the virus initially from China at the end of 2019, then from other hubs and cross-border in Europe, the task of AGES was that of a reliable, timely, and easily accessible information provider. As described above, the situation at the initial spread in Austria at the end of February 2020 was rather unclear in Austria. When the first two cases of COVID-19 were confirmed by positive tests on 25.02.2020, the knowledge of the stability/durability of the virus and feasible as well as sensible and effective protection measures were still not clearly described. Even first data sets regarding mortality were broadly varying between the different sources and health experts.

In parallel to the rapid spread of the novel Corona virus in Austria – see the equivalent of the newly reported cases per state in the nine Austrian federal states in the figure below– the urgent need for low-threshold, reliable and well-prepared information from the population developed.
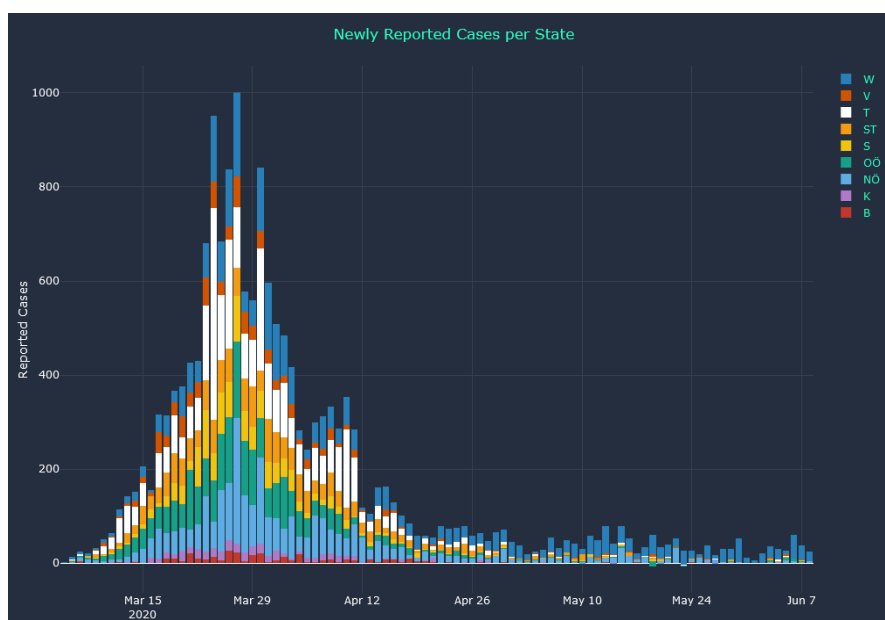
Figure 2: Confirmed cases change over time in Austria (CSH 2020)

The Austrian Agency for Health and Food Safety (AGES) is a company of the Republic of Austria, owned by the Austrian Federal Ministry of Labour, Social Affairs, Health and Consumer Protection and the Austrian Federal Ministry for Sustainability and Tourism. Founded on 1st June, 2002, AGES supports since nearly two decades the management of the federal ministries and the two federal agencies subordinated to it in questions relating to public health, animal health, food safety, medical and drug safety, food security and consumer protection along the food chain by providing professional and independent scientific expertise (as stated in Article 8 of the Austrian Health and Food Safety Act).

With its public service mandate from its owners, AGES was – together with other highly relevant tasks in the analytic sector – mandated as a central communication hub and responsible for knowledge distribution during the pandemic spread. From the spread of the CoV-2 in the neighbor countries of Austria – and specifically in Italy – the demand for reliable information and guidance was rapidly growing. As a first contact base, the CORONA hotline set up and hosted by AGES on very short notice, was one of the main means of communication and reliable distribution for the Austrian people from official side. The AGES hotline was soon expanded to a 24/7 schedule and supported by the regularly updated online service (https://www.ages.at/themen/krankheitserreger/coronavirus/#c77249) to give easily accessible information with a low-threshold. In addition, a new tool on twitter supports the tackling of fake news by stating on top of every *#coronavirus* query the link to the quality checked AGES portal. Although twitter is still not extensively used in Austria, this additional service supports the distribution of high-quality information and helps countering fake news.

Due to the high demand of the broadly-advertised hotline and the high quantity of lay persons not acquainted with medical and health care specifics, it is also evident, that a specifically tailored, approach of the call center experts was required to satisfy these information needs. Analogous to the following graphics, an orientation on the caller´s individual needs and the orientation on a transfer of highly scientific facts to digestible and broadly understandable contents was important for this service facility.
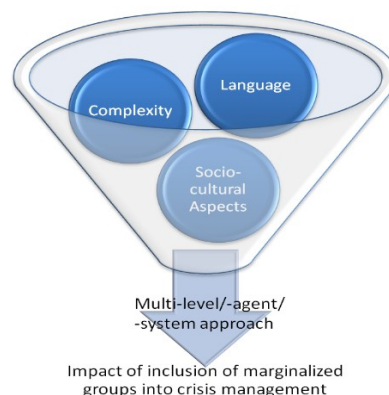
Figure 3: Requirements for high-impact, inclusive approach in crisis management (Rainer et al. 2016)

The quantity of information seekers and callers at the CORONA hotline was challenging the human and technical resources of the organization. Gaps in regards of the technicalities of the call center work in order to satisfying the demand were quickly resolved. However, it became obvious, that a certain amount of spare capacities – not only in regards of the health system per se but also in supporting management facilities – has to be kept in times between crisis. This has to be cautiously held upright to facilitating a timely set up of a high level, functioning and stable information hub to satisfy the urgent needs of the population for low threshold contact from public and thus reliable side.

Also, the contribution of AGES and its Department of Infectious Disease Epidemiology and Surveillance (INFE) to the Epidemiologic Reporting System (EMS), the database of all public health providers and the Federal Ministry of Health, was demanding a high logistic, technical, and human resources effort. Data visualized as the figure above shows, was updated several times daily in the rapidly evolving, dynamic pandemic situation to give the Austrian population as well as decision makers and stakeholders a timely and above all reliable, understandable view on the development.

Even if the information flows were pre-defined and outlined in crisis plans, the sheer mass of analytic work, data analysis and digestion, human and technological resources necessary were strained to the limit and show the potential gaps, that could arise from an even harder and more grievous hit of a crisis situation in this area.

3. Interoperability and Gap Analysis from a Research Perspective – AIT and the findings from a Project View

In 2016 Neubauer et al identified interoperability of stakeholders as a predominant requirement in crisis and disaster management. The results were achieved based on more than 50 systematic expert interviews with international disaster managers reporting on different type of disasters, such as flooding or wildfires. Figure 2 gives an overview on the type of the reported events.
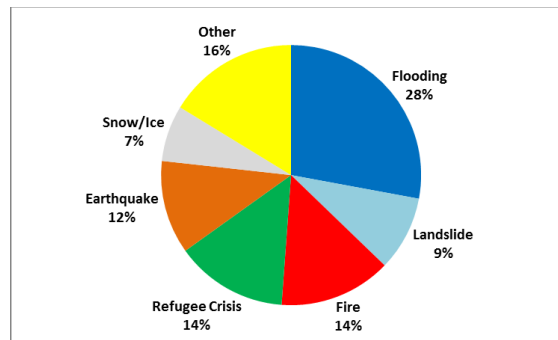
Figure 4: Type of reported disasters from the online survey from 2016 (Neubauer et al 2016)

In line with these results lack of interoperability was identified as a central capability gap in the frame of DRIVER+.

In more detail several facets need to be considered when discussing shortcomings related to interoperability:

- Challenges in exchanging relevant information between actors such as authorities and first responders

- Limitations in the common understanding of the exchanged information

- Problems in understanding differences in procedures of different organisations acting in the same scenario

- Lack of awareness of overall resources and lack of shared understanding

Due to these outcomes many demonstrators for interoperability were developed in research projects both on a national as well as an international level. Example given, within the European project EPISECC a Common Information Space (CIS) was developed demonstrating the feasibility of seamless message exchange between proprietary IT systems of stakeholders. The concepts of EPISECC were used to develop the CIS of the testbed of the large-scale European project DRIVER+. This CIS was used to test a multitude of crisis management solutions in the frame of five systematic trials demonstrating again the feasibility of syntactic interoperability.

The central idea of a CIS is to agree on a standardized protocol and to develop an adaptor for each of the systems connected to the CIS. This allows that each of the participating organisations can continue to use its legacy system as well as message format, the adaptor translates the messages to the common standard message format and allows smooth information exchange via the CIS. At the side of the recipient the message is again transferred to the message format of the system of the recipient. The advantage of a CIS becomes evident when comparing the option to interface each system with each other on a bilateral level versus connecting each system to a platform. As can be seen in Figure 3, the first approach requires the development of 20 interfaces in case of 5 systems, the use of a platform requests only the development of 5 adaptors. In case of systems bilateral interfacing requests 90 interfaces, a platform interconnecting 10 different systems needs only the development of 10 adaptors.
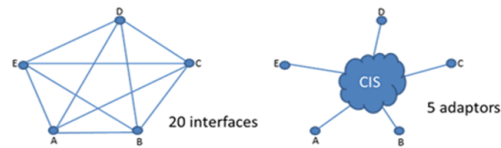
Figure 5: Bilateral interfacing versus the CIS concept (Neubauer et al 2017)

Within these projects interoperability of systems was focusing on information exchange between stakeholders while managing natural disasters such flooding, earthquakes or wildfires. The outcomes of such projects served as a basis for the development of market ready solutions such as the Public Safety Hub. The development of any interoperability platform always needs the fulfillment of some central requirements. An international working group decided therefore to standardize requests to realize information exchange in crisis and disaster management. The group specified the central needs to achieve interoperability on semantic and syntactic level. The outcomes of this work are published in the CEN Workshop Agreement (CWA) 17513. Some central requirements to realize interoperability are:

- Selection of an interoperability layer model for further development

- End to end transfer of information between stakeholders as well as fast translation of information from an operational point of view

- Agreement on a standardized protocol to route and transfer information

- A common catalogue of structured messages / data structure

- Verification of messages before routing them to receivers

- Syntactic mapping of the entities of the Command and Control systems of the information exchanging stakeholders to the interoperability platform data model

- Definition of a common terminology as a basis for semantic interoperability

- Establishment of a semantic mapping and matching process

Within DRIVER+ a repository of gaps of and solutions for the domain crisis and disaster management was developed and will be operated after the end of the project in order to give on one hand the practitioners the opportunity to specify gaps they are facing when managing different types of crisis as well as to identify solutions from industry, SMEs and research suitable to close such gaps. On the other hand, solution owner can present their solutions in the repository. The central strength of the platform is that both gaps as well as features are described using the same taxonomy making automatic matching of gaps and solutions possible. In the early stages of the spread of COVID 19 in Europe, it was decided within DRIVER+ to focus specifically on solutions serving to manage different facets of COVID 19 as well as gaps related to pandemic management. For this purpose, a questionnaire was developed and was distributed among international stakeholders in order to get a better understanding of the main challenges the different actors are facing. The outcome of these investigations is described in detail in Neubauer et al 2020, but a central finding is that very similar to the management of natural crisis and disasters interoperability turned out to be a cornerstone in pandemic management, too.

4. First Responders and Social Services in Response of the Crisis – lessons learned from Johanniter Unfallhilfe Austria

Johanniter Unfall-Hilfe in Austria is an ambulance and care provider, which is active since 1974 across Austria. Within the last 10 years, they grew especially in Vienna and widened their activities

from ambulance and patient transport to mobile care services, personal emergency alert systems, disaster relief, education and training of health professionals and research. During CoV-19, Johanniter introduced a crisis staff and Liaison Officer following the national standards SKKM (which is related to the INSARAG guidelines) to support information exchange and implementation of safety measures for staff and patients in the health care sector.

Johanniter are also active in several research projects related to active and healthy ageing and security research. This chapter shall provide a short introduction to the use cases related to CoV-19 from the view of Johanniter.

4.1. CoV-19 impact on ambulance service and patient transport

Cov-19 had a strong impact on the ambulance and patient transport system in Austria. Hospitals aimed at optimizing the capacities for CoV-19 patients. Therefore, everyone who was not in a critical state was evaluated if they could be sent back home. Non-urgent surgeries were postponed in order to reducing the capacity utilization of intensive care units. Due to this, in the first days of counter-CoV-19 measures in Austria, patient transports stayed at a high level to transfer patients home from hospital care. After this first period of action, a massive drop in patient transport was noticeable.

For the ambulance and patient transport staff, increased hygiene rules were established, and additional personal protective equipment was used. The usage of FFP2 masks became a standard during the first weeks. Due to a lack of supplies, this was changed to wearing surgery masks for standard transports and in case of signs of CoV-19, FFP2 masks were used. In addition, a light CBRN-protection suit had to be worn by the ambulance staff.

The increased use of personal protective equipment was complemented by using disinfection fluids. Here as well, the massive demand in the population led to additional problems in supplying ambulance and patient transport with sufficient supplies. The supplier of Johanniter in Austria started to produce disinfection fluids using recipes provided by the World Health Organization and the support of pharmacies.

To coordinate the cooperation with other elements of health care and government, each organization started to set up an emergency relief team. In Austria, the standard of the national SKKM (State Crisis and Disaster relief management, BMI 2020) was the baseline for actions. This standard is aligned with the German DV100 and the INSARAG guidelines for international operations.

These emergency relief teams, or crisis staffs were linking the health care organizations with the government interface. This allowed a centralized procurement of supplies and a distribution by the government towards the organizations according to the needs to keep up the functional capability. The central crisis staff of the Ministry of the Interior and the Ministry of Health updated the ambulance and patient transport organizations in Austria on a daily base. This was conceptualized as a two-way flow of communication. All health organizations reported back to the government about CoV-19 related issues and potential infection rates.

During the time of CoV-19, at Johanniter in Austria, there was no indication of a higher infection rate or sick leave rate in ambulance staff and patient transport staff in comparison to years before.

4.2. CoV-19 and its impact on research around active and healthy ageing and ambient assisted living

When CoV-19 became a problem in Europe, projects for active and healthy ageing (AHA) and ambient assisted living (AAL) were affected as well. To restrictions in personal contact and safety measures for high risk population made it nearly impossible to roll out technology for AAL and to deliver interventions for AHA. Next to potential delays in tests, this was also disrupting study designs

like Randomized Controlled Trial studies or other experimental approaches. Both types of projects, AHA as well as AAL, aim for a target group 60+. Most of the participants were invited to studies because of additional health related limitations. In the case of the project My-AHA, these were people with clinical frailty. In the case of the project 24hQuAALity, participants were chosen because of the need for a 24h home help.

The CoV-19 situation was pushing digitalization in the health care sector to an extend that was not known until then. Within several weeks, the health sector had to react to a new condition and implement solutions for the contact problems as fast as possible. In Austria, this led to a fast uptake of medication prescriptions at the physicians. What was planned by the electronic health record agency for the upcoming years was implemented within days. Physicians were allowed and encouraged to send standard medications for long term treatment directly to the pharmacy of the patient. The patient could pick up prescribed medication at the pharmacy by showing the health insurance card (in Austria: eCard). For AAL, this is a starting point in the direction of online prescriptions.

In AAL projects concerning personal emergency alert systems, the research department at Johanniter in Austria experienced an increasing request for this kind of products. Nevertheless, due to the development state of the AAL projects, people send devices back after contact restrictions were loosened.

In AHA projects, contact restrictions led to a problem in the delivery of physical trainings and social meetings. Thereby the efficiency of interventions was disturbed. To reduce this negative effect on studies and participants, personal trainings like the OTAGO program were recommended to allow a self-organized training. In addition, the contact systems and messengers were of more importance to the participants. As these systems work with limited functions in the projects, no effects were measured.

Another issue in both project types that aim to increase the quality of life, is the impact on the psychological wellbeing of the target group.

4.3. CoV-19 and its impact on Security Research

CoV-19 touched many aspects of life and business and thus, security research was also affected. With the project STAMINA (Project no. 883441), a disease control system should be developed. The special use case for Johanniter is the spread control in patient transport and ambulance service. The project was submitted and granted before CoV-19 but became more important to the situation at hand. During the CoV-19 situation, a lot of data material was collected in advance to support this new project, starting in September 2020, long after the first wave of CoV-19.

Another aspect relevant for security research could be the situation on cruise ships. In the project PALAEMON (Project no. 814962), security systems for evacuation of cruise ships are developed and a Mass Evacuation Vessel is designed and to be tested. Due to the CoV-19 situation, the basic problem was turned around. How can people stay safely in their cabin? How is it possible to run disease control on a cruise ship? These new questions are actively discussed as well as the question what happens if a cruise ship with 5000 passengers has to be evacuated while there is CoV-19 on board? Also tracking devices on board could support an estimation for disease spread and support in the choice of the right countermeasures.

Another important aspect of the CoV-19 crisis was the monitoring of the disease in refugee camps at the Greek-Turkish border. Just before the start of CoV-19, rumors of tens of thousands of refugees starting to cross the border between Turkey and Greece flashed up and reminded at the situation of 2012. Fired up by speeches of the Turkish President with the intention to open the borders for refugees, it was of utmost importance to start monitoring the situation and collect as much

information as possible to prepare for incoming refugees within Europe. Johanniter were part of the national funded project HUMAN+, a bi-lateral action between Austria and Germany to support preparation for refugee wave in central Europe. Therefore, social media data was collected and monitored (Twitter). When the Turkish government started rumors about opening borders, HUMAN+ was put to action. Very clearly, it turned out to be a rumor and not tens of thousands were moving towards Greece but 5.000 or less.
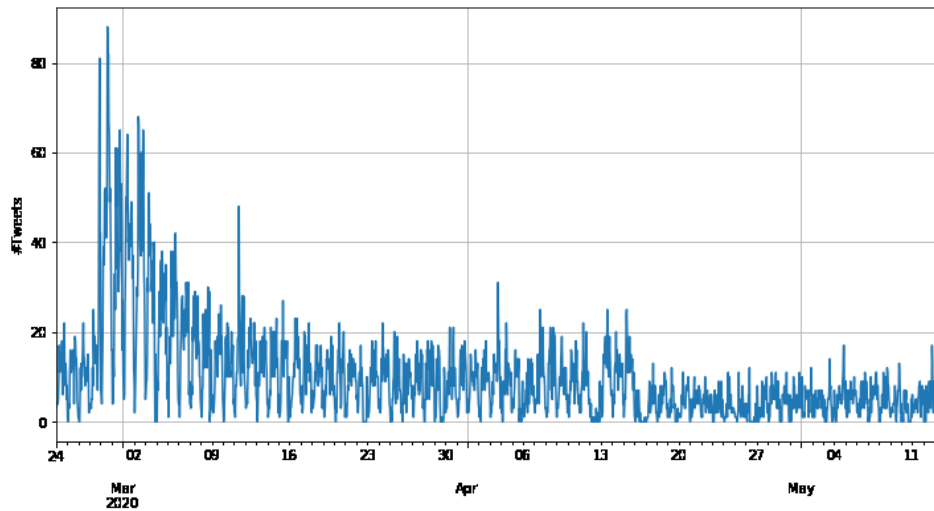
Figure 5 HUMAN+ Data on refugees 2020

This data showed a fast raise of social media in a short time but also a fast decline within a couple of days. The interpretation was that the pressure at the borders were just a short raid and not a permanent situation. This indicated the state of rumors and not for open borders. It could be also shown that refugees moved northwards to cross other borders as well. Which allowed to prepare support for different routes.

With this first promising result, when CoV-19 started, a major concern was the situation in the refugee camps. A social media monitoring for CoV-19 and flu related tweets was started as well. The data was also showing that the situation in camps could be interpreted as serious but stable. At the critical time between 10th March 2020 and 10th April 2020, a high rate of activity could be seen, but with a tendency to slow down. Since mid of April, the situation seemed to be under control and slightly decreasing.
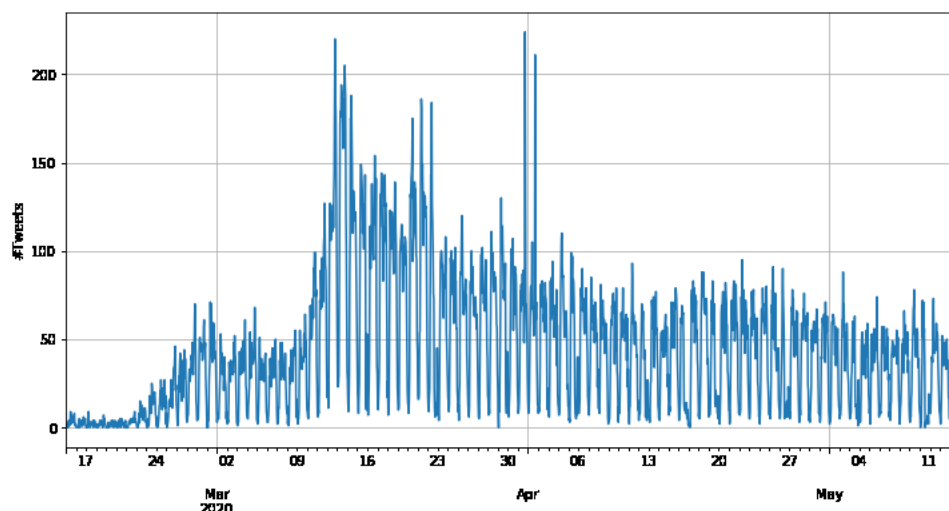
Figure 7 CoV-19 data HUMAN+, March 2020 - May 2020

161

This data, provided by Paris Lodron University Salzburg, Dpt. Z_GIS, and their spin off Spatial Services Ltd., allowed planning and preparations according to actual needs. The HUMAN+ system was used in Germany by the Bavarian Red Cross as well as in Austria by Johanniter for coordinating actions and resources management. In this special case: realistic estimation of a calm situation. Also, for the estimation of a potential 2nd wave of CoV-19, this data was helpful until now to evaluate the situation and support decision making processes.

5. Summary, outlook and approaches for further discussion

The situation after half a year of the initially rather diffuse, but quickly and dynamically evolving CoV-2 spread in Austria can be evaluated as relatively positive despite the highly discussed and not un-criticized measures of the early lock down. Its partially pressing societal, economic, and also critical psychological side effects and impacts (see e.g. Metzler et al. 2020) and basically the necessity of some measures is highly controversial in scientific but also in broad circles.

This exemplary overview of some unique use cases with Austrian connex showed how it was not only possible to "flatten the curve" of new infections. Several factors are still very unclear and will never be conclusively answered (e.g. how models of infection are correctly set up s. Czepel 2020), even for a relatively well researched and documented event as the CoV-2 spread. However, it became evident, that several factors were up to now and will in a future, possibly even more severe health crisis with higher mortality, morbidity, and risk of contagion, be of utmost relevance for the management and above all the minimization of a spread and lethal effects. Several factors that are partially common knowledge, best practice, but often also open issues in crisis management shall be stated briefly for further discussion and as a potential kick off for future research:

One of the major factors is – as also in traditional crisis and disaster management – related to *information and its management* (Rainer et al. 2015 and 2016). Timely information in a digested, *clearly visualized* manner, verified data, the availability of selected and well researched sources on an international level is of vital importance for the sensible and solidly based decision-making process in crisis situations.

Open source data and free exchange of high-quality data like the good practice example see e.g. Open Data Österreich and other services offering their usually high cost services for free during the hot phase of the pandemic spread show the possibility of scientific collaboration in times of need and a common drive to tackle a critical situation.

Political will and cooperation were key for the timely actions that were – if not undisputed and for sure not without flaws in Austria – overall to be rated successful in containing a wide spread in a first wave (see e.g. CSH 2020).

Timely communication management and provision of *validated information* is and was crucial for compliance, coordination, and adaptive measures in management to tackle dynamic situations as the COVID-19 pandemic showed. *Low-threshold information services* like the CORONA hotline and online services contributed to satisfying this pressing requirement of managing the crisis in Austria.

Credibility of information sources for the public, also reflected in the validated information work of AGES and other specific services were thus vital for the success of public measures.

In the same train of action, *challenging of fake news* and the timely stop of spreading e.g. via the twitter #coronavirus initiative of AGES leading to quality checked, reliable information.

After all, also the maintaining of resources for crisis management – hard-, soft-ware and human resources – became evident to being a key factor for successful management and a resilient society.

In addition, the all-time issue of interoperability in and between service providers was a crucial issue that had to be taken into consideration at all times in this as well as in any crisis situation.

Use and analysis ability for *multiple data and sensor sources* (see Rainer et al. 2017) for e.g. inclusion of heat sensors for mass screenings could be one possible way to tackle future health crises with even broader impact on the health system.

A common language and data requirements inter organizations is obviously a still pressing need for the successful management and tackling of emerging events on a general basis.

The same applies for multi-use approaches e.g. for logistic and communication issues, but also for the use of devices and production chains like in the case of 3D printing of respiratory systems.

Also, the relevance of directories and stock taking for critical logistics and goods is not only applicable for the current situation but for an overall set up of preparedness and resilience vis-à-vis dynamic and complex crises and disasters. Thus, network, exchange, and interoperability issues between different stakeholders on various levels (local/regional/national/cross-border/international; different organizational backgrounds; public/private…) are key for the management of such events.

Last but not least, feasible and low-threshold approaches for public information, compliance raising and population acceptability of measures as elaborated in the use cases is another key factor for a supportive and sustainable crisis management.

Overall, CoV-19 was a challenge that had to be conquered in many ways. The flexibility of organizations has been tested and potential for development could be identified in its wake to increase the preparedness in the future. For emergency first responders, working under the situation of a potential infection is technically a standard situation. However, at the beginning of the outbreak, too many uncertainties steered up the chaos phase but even with a high risk of exposure, cases of CoV-19 were very rare in Ambulance Services. If this was a lucky coincidence or not could be element of further research.

The impact on research in different areas was clearly a delaying factor but also providing chances to react to urgent needs. It was shown that research can and should be shaped and focused to presenting a strong support to fighting pandemic spread and the coordinated, timely and least-disruptive measures of containment.

6. **References**

Biswaranjan, P., Kabita, D., & Sarat Kumar, P. (2020). Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India. Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India. In: Science of The Total Environment, Volume 728, 2020. 138914, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.138914. (http://www.sciencedirect.com/science/article/pii/S0048969720324311)

Borges do Nascimento, I .J., Cacic, N., Abdulazeem, H. M., von Groote, T.C., Jayarajah, U., Weerasekara, I., Esfahani, M. A., Civile, V. T., Marusic, A., Jeroncic, A., Carvas Junior, N., Pericic, T. P., Zakarija-Grkovic, I., Meirelles Guimarães, S. M., Luigi Bragazzi, N., Bjorklund, M., Sofi-Mahmudi, A., Altujjar, M., Tian, M., Arcani, D. M. C., O'Mathúna, D. P. & Marcolino, M. S. (2020). Novel Coronavirus Infection (COVID-19) in Humans: A Scoping Review and Meta-Analysis. J. Clin. Med., 9, 941.

BMI Federal Ministry of the Interior (2020). Civil Protection in Austria. National Crisis and Disaster Management (SKKM). https://www.bmi.gv.at/204_english/skkm/start.aspx (retrieved 19.06.2020)

BMSGK Federal Ministry of Social Affairs, Health, Care and Consumer Protection (2020). Amtliches Dashboard COVID 19 https://info.gesundheitsministerium.at/ (retrieved 22.05.2020)

CEN Workshop Agreement 17513, Crisis and disaster management – Semantic and syntactic interoperability, May 2020, https://www.cencenelec.eu/research/CWA/Pages/default.aspx (retrieved 11.06.2020)

CSH Complexity Science Hub Vienna (2020). https://www.csh.ac.at/coronavirus-how-measures-work-wirksamkeit-der-massnahmen-oesterreich-april/ (retrieved 04.06.2020)

Czepel, R. (2020). COVID-19. Streit hinter den Kulissen. Science.ORF.at https://science.orf.at/stories/3200563/ (retrieved 07.06.2020)

Gianicolo E., Riccetti N., Blettner M. & Karch A. (2020) Epidemiological measures in the context of the COVID-19 pandemic. Dtsch Arztebl Int 2020; 117: 336–42.

Ludwig, S., & Zarbock, A. (2020). Coronaviruses and SARS-CoV-2: A Brief Overview. Anesthesia and analgesia, 10.1213/ANE.0000000000004845. Advance online publication. https://doi.org/10.1213/ANE.0000000000004845

Mackenzie, D. (2020). Covid-19: Why won´t the WHO officially declare a coronavirus pandemic? In: NewScientist. 26.02.2020, updated 11.03.2020. https://www.newscientist.com/article/2235342-covid-19-why-wont-the-who-officially-declare-a-coronavirus-pandemic/ (retrieved 03.06.2020)

Metzler, H., Rimé, B. & Garcia, B. (2020). Expressions of emotions and social behavior around COVID19 on Twitter in different countries. https://csh.ac.at/covid19/twitter/covid_twitter_Austria.html (retrieved 05.06.2020)

Neubauer, G., Preinerstorfer, A., Schirnhofer, S., Humer, H., Lichtenegger G., Vorraber W., Linke H., Tusa G., Gruener R., Dalaff C., Knezic S., & Blaha M. (2016). Validation of the Management of Past Crisis and Disasters In: Petr Doucek et al. (Eds.). IDIMT 2016. Information Technology, Society and Economy Strategic Cross-Influences. 24th interdisciplinary Information Management Talks. Trauner Linz. 123-130.

Neubauer, G., Preinerstorfer ,A., Lichtenegger, G., Humer, H., Linke, H., Zuba, G., Delprato, U., Tusa, G., Knezic, S., Baucic, M., Dalaff, C., Gurczik, G., Ammar, B. & Magzoub, A. (2017). Common information space as enabler for collaboration in disaster management: Demonstration of the validity of the EPISECC CIS concept. In: Proceedings of ICT-DM 2017, International Conference on Information and Communication Technologies for Disaster Management, IEEE, (2017), ISBN: 978-1-5386-2553-8; Paper-Nr. 18, 8 S.

Neubauer G., Preinerstorfer A., Martin A., Rainer K., & van Berlo M. (2020). Analysis of Gaps Arising while Managing the COVID 19 Crisis: Submitted to IDIMT – Interdisciplinary Information Management Talks, Czech Republic, Kutna Hora, September 2nd- 4th, 2020.

Open Data Österreich. www.data.gv.at/covid-19/ (retrieved 05.06.2020)

Rainer, K., Levy, I., Schmid, J., Götsch, K., Quirchmayr, G., Göllner, J., Forst, N., & Backfried, G. (2015). New opportunities and challenges for participation in crisis and disaster relief. The QuOIMA project as example for interaction, participation and privacy protection in disaster management. In: Central and Eastern European Legal Studies I/2015. 25-41.

Rainer, K., Neubauer, G., Ruzsanyi, V., Silvestru, D., Almer, A. & Lampoltshammer, T. (2017). The Potential of Multiple Types of Sensor Data and Information Exchange. Challenges and Perspectives for an Operational Picture for the Response to Crises with Mass Involvement. In: Petr Doucek et al. (Eds.). IDIMT 2017. Information Technology, Society and Economy Strategic Cross-Influences. 25th interdisciplinary Information Management Talks. Trauner Linz. 111-126.

Rainer, K., Neubauer, G., Pointner, K., & Kastel, W. (2016). Inclusion of Marginalized Groups into Crisis Management Processes. Models, Case Studies, and holistic Perspectives. In: Petr Doucek et al. (Eds.). IDIMT 2016. Information Technology, Society and Economy Strategic CrossInfluences. 24th interdisciplinary Information Management Talks. Trauner Linz. 87-100.

Roser M., Ritchie, H, Ortiz-Ospina E., & Hasell J. (2020). "Coronavirus Pandemic (COVID-19)". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/coronavirus' (retrieved 04.06.2020)

SailLabs (2020). Media Mining Client for Media Monitoring of Pandemic Outbreak – Coronavirus. Vienna.

Vaezi, A., & Javanmard, S. H. (2020). Infodemic and Risk Communication in the Era of CoV-19. Advanced biomedical research, 9, 10. https://doi.org/10.4103/abr.abr_47_20

W.H.O Rolling updates on coronavirus disease (COVID-19) (2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. (retrieved 10.04.2020)

ANALYSIS OF GAPS ARISING WHILE MANAGING THE COVID-19 CRISIS

Georg Neubauer, Alexander Preinerstorfer, Andreas Martin

AIT Austrian Institute of Technology GmbHgeorg.neubauer@ait.ac.at, alexander.preinerstorfer@ait.ac.at, andreas.martin@ait.ac.at

Karin Rainer

AGES Austrian Agency for Health and Food Safety
karin.rainer@ages.at

Marcel van Berlo

TNO The Netherlands Organisation for Applied Scientific Research
marcel.vanberlo@tno.nl

Abstract

In the frame of the research for DRIVER+ strategic gaps arising in the management of crisis and disasters were analysed and used as basis for the development of the products and methods of the project. To reflect the requirements arising in the current COVID-19 crisis it was decided to search for challenges and gaps related to the management of pandemics. For this purpose, an electronic questionnaire was developed and distributed using the networks of the project. In the frame of this paper we present the results of the analysis of the first stage of distribution reflecting the situation by the end of April 2020. The results demonstrate a wide range of existing challenges, encompassing lack of different type of resources as well as interoperability and data management gaps.

1. Introduction

In 2020 the COVID-19 crisis is the predominant challenge of the world and will most likely continue to have considerable impact on our society in the upcoming years. However, from an historical point of view this pandemic is not unique at all. E.g., the black death (bubonic plague) taking place from 1347 to 1351 caused estimated 200 million fatalities. In more recent times the Spanish Flu taking place from 1918 to 1920 was a pandemic causing the death of estimated 40 to 50 million people. Aids started to spread around 1981 and led to death toll of 25 to 35 million persons. Looking at the last 20 years SARS (2002-2003, 770 fatalities), the Swine Flu (2009-2010, 200.000 fatalities), MERS (starting in 2012, 850 fatalities) and Ebola (2014-2016, 11.300 fatalities) are the most prominent

recent pandemics (Visual Capitalist, 2020). The mortality is of course only one suitable indicator to compare different pandemic events, among the multitude of other parameters to be considered are the morbidity or the economic impact. Placing the ongoing COVID-19 crisis in this context, 4.2 million cases and 287.000 global deaths are reported mid May 2020 (COVID-19 Dashboard, 2020). Morbidity and mortality rates are strongly varying across the different affected countries, e.g. an incident rate of 0,36% is reported at this stage from Italy, whereas 0,07% are reported from Slovenia. When comparing the different statistical data from different countries one must be aware that different approaches for testing and counting infected persons as well as assigning of reasons of deaths are applied making comparisons between countries or even regions difficult. It has to be noticed that these data do not take indirect effects arising to the cancelation of medical treatments or reluctance of patients to go to hospitals into account. In order to start to systematically identify and analyze gaps and requests arising in the course of the management of COVID-19, the project team of DRIVER+ developed and distributed a questionnaire on the different types of challenges in the management of COVID-19. This paper presents the outcome of the first 55 responses to this questionnaire. The outcomes of this inventory are used to earn background information about predominant gaps as well as emerging solutions to have a better basis to populate the Portfolio of Solutions of DRIVER+ (PoS, 2020). This investigation encompassed the collection of information on already available as well as requests on potential solutions that are suitable or needed to close some of the identified gaps. The strategic target of this initiative is to provide a basis for an improved management of pandemic events in the future by analyzing these gaps and solutions.

2. Method

The questionnaire on COVID-19 was implemented using the Survey Monkey Toolkit (2020). Following an iterative approach, the first version of the questionnaire went online in English on 08.04.2020. The adaptive presentation logic allows the questionnaire to be accessed on desktop devices as well as on mobile devices.

In the first dissemination phase, the questionnaire was send to the external cooperation network of the project as well as to the Centers of Expertise of DRIVER+, the research and technology organisations that are members of the EARTO Security & Defence Working Group and all members of the European Organisation of Security, as well as to the project partners of the uprising H2020 project STAMINA leading to a total number of contacted persons of approximately 590. The survey was also shared via DRIVER+ LinkedIn and Twitter, the LinkedIn message was shared by DG ECHO and JRC. The addresses were directed to the survey's welcome page via hyperlink (Collector 1), where objectives and background of the survey are described. First background information on the responding professional and the organization he is belonging is requested. On the following parts of the survey the specific gaps are queried. For each central gap related to COVID-19 a multiple-choice question is provided. Additionally, each gap offers the possibility of a free text answer. Finally, a free text area is given for a short description of existing or requested solutions to solve specific gaps. To emphasize the voluntary nature of the answers and to positively influence the drop-out rate, only questions about the professional background of the respondent are mandatory. In total the questionnaire consists of 12 pages with 23 questions. In a second iteration both the questions and the answers were peer-reviewed and improved, to further increase participation the new version of the survey was translated into German, Estonian, French, Italian, Dutch, Polish, Swedish and Spanish in addition to the initial English version. This version went online on 13.05.2020. The third iteration is based on targeted adaptations that were made by analyzing the preliminary results. Corresponding adaptations were carried out online and the adapted version of the survey was sent out with a 2nd collector on the 26.05.2020.

By the end of May 2020 99 international respondents provided information via the online questionnaire leading to an estimated response rate of about 17%, the level of information provided (e.g. number of questions answered) varied considerable between the respondents. Regarding the response rate addresses were invited to distribute the questionnaire to other experts in the field. The number of these "secondary" contacted persons is unknown to the authors, therefore the real response rate is below the indicated 17%. The personal data collected in this survey are used for research purposes only. To categorize the pandemic requirements a taxonomy of gaps was developed. The main sections of the questionnaire are corresponding to the main gap categories of the taxonomy, namely Communication and Information Exchange, Data Management, Resource Planning and Logistics, Cooperation with and Management of the Population, Safety and Security and Public Health. The respondents have the option to select only gaps belonging to a subset of categories and are therefore not forced to answer all questions.

3. Results

In this chapter the central outcomes of the first dissemination phase to the questionnaire are presented. These answers were received within two weeks after the first distribution of the questionnaire and reflect the view of the responding experts by the end of April 2020. During the first wave of investigation 55 responses from 21 countries were received. 46 of the respondents are from Europe. Regarding the profession of the respondents the largest groups were working in the health and the research area as well as crisis managers. Each of these subgroups are about 20 % of the respondents providing information on their profession.

The category of gaps "Communication and Information Exchange" was selected by 31 % of the respondents. Looking at the specific gaps, challenges related to "The Role of Social Media in Your Organisation" and "Governance and Related Actors" were the most frequently selected gaps as can be seen in Figure 1. Other frequently selected gaps were "Lack of information distribution, not harmonized data format" as well as "Lack of a Common Operational Picture". The analysis of the free text areas allowed to specify other gaps, challenges related to task distributions and coordination were reported in the free text areas. One respondent reported on problems related to both communication and data management. Among other problems, the challenges arising due to the different standards for reporting the deaths due to COVID-19 at European level as well as the uncertainty about the officially reported values regarding SARS-CoV-2 infections due to the very limited number of tests were stressed. Not in line with the other gap specifications, one respondent stated that communication and information exchange are working very well, another one reported about growing communication capacity indicating thereby an improvement of the communication capabilities in the timely course of the COVID-19 crisis. 27 % of the respondent identified gaps related to resource planning and logistic. It turned at that the predominant challenge is related to lack of resources and materials, other less frequently stressed gaps are related to "lack of cross-organisation overview of resources (human as well as material)" and "lack of knowledge on available production capabilities". The lack of personal protection equipment is mentioned as well as lack of an EU wide system to share and use available expertise and resources. More generally, insufficient stocks of equipment, consumables and protection at hospitals and population care centers is stressed by one respondent. He is also outlining a non-transparent policy for the evolution of stocks. Looking at the other categories of gaps, a multitude of additional needs became visible. The category "Impact Assessment" was selected by 27% of the respondents. Lack or limited capabilities to monitor risks turned out to be the predominant gap category, followed by lack or limitations related to contingency plans. One responder claimed an excess of non-specialists performing epidemiological modelling, another expressed concerns because of a lack of understanding of the economic consequences of medical measures. Concerns about gaps related to data management was expressed by 20 % of the

respondents, the most prominent gap is tracking and tracing of suspected cases. In one case concern was expressed regarding the officially reported infected cases due to the low number of tests, another respondent was reluctant regarding the suitability of so far applied technical solutions that made matters worse. 24 % of the respondents selected one or more gaps belonging to the category "Cooperation with and Management of the Population". Predominant gaps in this category are problems in reaching parts of the population, education, training and risk awareness and challenges related to the management of fake news. The impossibility for voluntary testing of parts of the population being interested in their health status was stressed. In a more generic way, a distinction between system and capacity gaps was expressed. One respondent expressed his opinion in stating that the majority of the problems are related to human and not to system factors.

Safety and security related gaps were selected by 24 % of the respondents. The by far predominant gap of this category is related to the management of non-cooperative parts of the population, followed by limitations of available means to help affected parts of the population and a lack of a database of extreme cases. Interestingly, the category of public health issues was only selected by 18 % of the population. The most frequently selected gap of this category is the lack of standardized tests followed by challenges related to training and education as well as triage. One respondent explicitly stressed the lack of rapid standardized tests. Another respondent expressed the suspicion that more people might decease due to the removal of standing health care arrangements than from the CoV-2 itself. According to this statement, the capacity to manage the COVID-19 crisis might cause cessation or limitation of normal public health services causing even higher costs than the COVID-19 crisis.

In addition to this snapshot on the experiences from a larger community of crisis managers, some further aspects on the gap analysis can be deduced from the involvement of AGES, the Agency of Health and Food Safety in Austria. As a public service provider with the mandate from the Federal Ministries, AGES was among other highly relevant tasks responsible for the central steering of the communication and knowledge distribution in the hot phase of the crisis. From the spread of the CoV-2 in the neighbor countries of Austria – and specifically in Italy – the demand for information and reliable support was rapidly growing up to an overwhelming level, challenging the human and technical resources of the organization. Gaps in regards of the technicalities of the call center work to satisfying the demand were quickly resolved. However, it became obvious, that a certain amount of spare capacities – not only in regards of the health system per se, but also in supporting management facilities – has to be kept in times between crisis. This must be cautiously held upright to facilitating a timely set up of a high level, functioning and stable information hub to satisfy the urgent needs of the population for low threshold contact from public and thus reliable side.
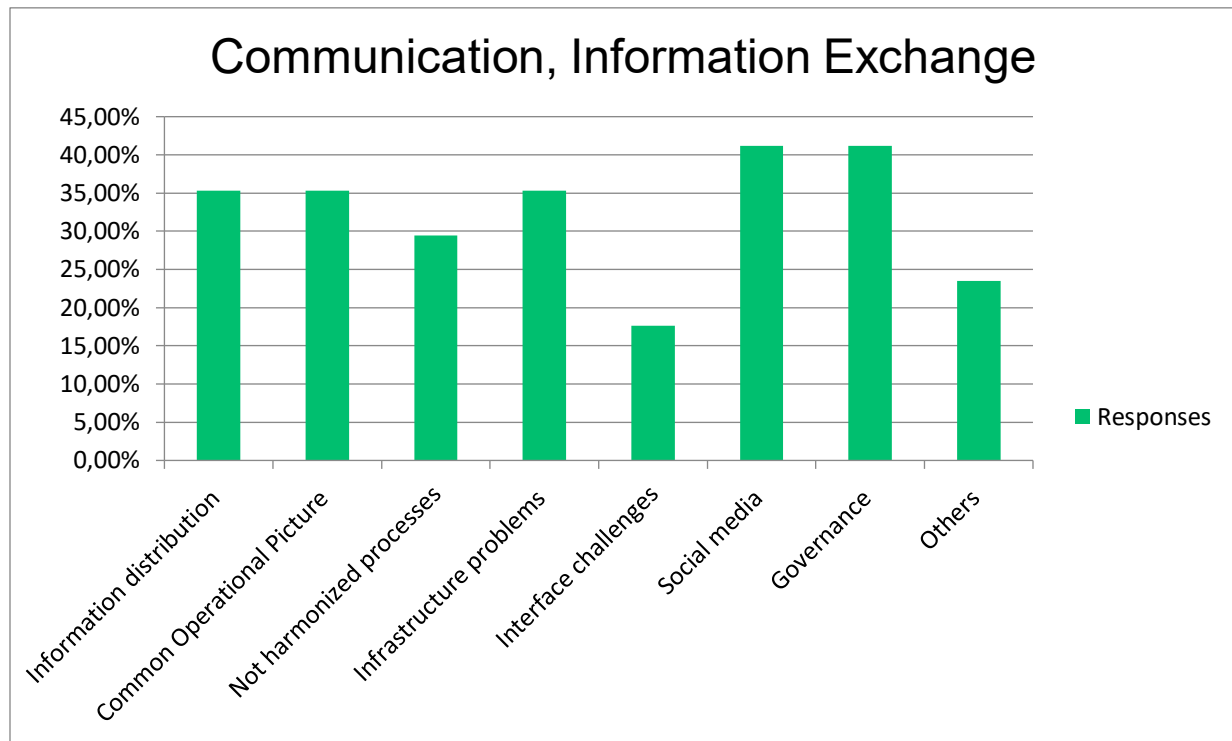
Figure 1: Results on the gap category Communication and Information Exchange

As a first contact base, the CORONA Hotline set up and hosted by AGES on very short notice, was one of the main means of communication and reliable distribution for the Austrian people from official side. The rapid development of the dynamic crisis regarding the regional spread and finally also the first wave of the pandemic infection in Austria was accompanied by a massive and rapid rise of information need – also via the AGES hotline and the online service (AGES, 2020). An incredible wave of information seekers contacted the well-advertised hotline that was soon expanded to a 24/7 schedule. Also, the contribution of AGES and its Department of Infectious Disease Epidemiology and Surveillance (INFE) to the Epidemiologic Reporting System (EMS), the database of all public health providers and the Federal Ministry of Health, was demanding a high logistic, technical, and human resources effort. Data visualized on a public dashboard was updated several times daily in the rapidly evolving, dynamic pandemic situation to give the Austrian population as well as decision makers and stakeholders a timely and above all reliable, understandable view on the development BMSGK, 2020). Even if the information flows were pre-defined and outlined in crisis plans, the sheer mass of analytic work, data analysis and digestion, human and technological resources necessary were strained to the limit and show the potential gaps, that could arise from an even harder and more grievous hit of a crisis situation in this area.

4. Conclusion and Outlook

The first version of the questionnaire that was used to obtain the results presented in this paper has some limitations. First, the information on the function of the respondents in pandemic management and civil protection of the respondents was unstructured. This is mainly caused by the free text area provided to answer this question instead of a structured multiple-choice selection. Another limitation is the so far low number of respondents from the same country. Out of 55 responses received 35 indicated their country of origin. The 35 respondents are from 24 different countries, in only 3 cases three or four respondents were from the same country. In the majority of cases only one respondent per country answered the questionnaire. It can be expected that the challenges in managing the

pandemic crises in the various countries vary at least partially considerably. One reason for this assumption is the different dimension the countries are affected by the COVID-19 crisis. An indicator for this is the differences in the incident rates, e.g. in mid-May 2020 an incident rate of 0,345 was reported for the UK, whereas the incident rate for Greece was only 0,027. The degree of contamination is of course only one indicator. There are many other factors that lead to different situations, such as differences of levels of preparation, availability of medical resources or different crisis management processes. These factors limit the value of combined comparisons of responses from different countries, but they give a good impression on the range of problems the world is currently facing. The results of this questionnaire give an insight on the type of challenges triggered by SARS CoV 2 by the end of April 2020. The result of this questionnaire is going to be used as a basis for a more in-depth analysis of COVID-19 related challenges to improve pandemic management in the upcoming years. The questionnaire is based on very straightforward, simple data model of pandemic gaps and does not replace future statistical evaluations of COVID-19 gaps.

5. Acknowledgment

6. References

AGES, https://www.ages.at/themen/krankheitserreger/coronavirus/#77249, (retrieved 29.05.2020)

BMSGK Federal Ministry of Social Affairs, Health, Care and Consumer Protection (2020). Amtliches Dashboard COVID-19, https://info.gesundheitsministerium.at/, (retrieved 22.05.2020)

COVID-19 Dashboard, John Hopkins University, https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6, (retrieved 29.05.2020)

PoS, DRIVER+ Portfolio of Solutions, http://pos.driver-project.eu, (retrieved 29.05.2020)

SURVEY Monkey Toolkit, https://www.surveymonkey.com/, (retrieved 29.05.2020)

Visual Capitalist, https://www.visualcapitalist.com/history-of-pandemics-deadliest/ , (retrieved 29.05.2020)

MULTI-MODAL AND MULTI-SENSOR APPROACH TO SUPPORT SECURITY MANAGEMENT TASKS

Alexander Almer, Thomas Schnabel, Anna Weber, Armin Köfler, Roland Perko, Stefan Ladstätter, Manfred Klopschitz

JOANNEUM RESEARCH Forschungsgesellschaft mbH
{firstname.lastname}@joanneum.at

Keywords

Abstract

Recent changes in the dynamic characteristics of security and emergency management requirements have a big impact on the development of more effective security systems. Innovative solutions have to provide support for new security challenges and newly developed operational strategies. Widely used security concepts based on video sensor solutions are being replaced by high-performance multi-sensor systems, which allow monitoring of outdoor areas with various sizes independent of lighting and weather conditions. Apart from the sensors used, the system effectiveness depends on innovative automatic data analysis including AI methods to enable generation of a comprehensive near real-time representation of a given situation.

This work describes a concept and an experimental prototype of a self-sufficient, independent multi-sensor system for monitoring of critical personal behavior in the context of events, border area surveillance and critical infrastructure monitoring. The specific security scenarios were defined and worked out in detail in close cooperation with security experts. In order to ensure 24/7 operational capability, the work focuses on a flexible sensor platform concept and includes optical, thermal and radar sensors which are synergistically employed. Furthermore, innovative approaches to multi-sensor AI-based data analysis and web-based mobile information management solutions were developed as part of the nationally funded research project WatchDog to enable a targeted, scenario oriented support of the security forces.

1. Introduction

The recent past has shown that dynamically changing security situations as well as required increase in efficiency in human resources raise the demands on emergency and security services as well as police and military forces. In the past few years, these problems have become apparent particularly in frontier operations in the context of the international migration situation. Likewise, the ever-increasing security requirements in industrial facilities and at large events place dynamic requirements on security solutions and management. These increasing demands on security forces require directed assistance with innovative technical solutions to realize an optimized support of time critical decision processes. To perceive threats, suspicious persons and critical behavior of crowds, information products from different sources (information systems, sensors, exploitation systems)

have to be available. The data has to be accessible with respect to time and location of the product in addition to other decision-relevant information, which should be incorporated as well. Integrated systems achieve enhanced situation awareness by developing a common picture of the area of interest. To support analysts, operators and decision makers it is important to integrate the correct, that is, temporally relevant and comprehensive information in this common situational picture in a user-friendly manner (Essendorfer et al., 2016).

In Almer et al. (2016), a comprehensive monitoring of movement behaviour and raising dynamics in crowds was described, which enables an early detection and prediction of critical situations that may arise at large-scale events. The paper presents a video-based airborne monitoring system enabling the automated analysis of crowd dynamics and to derive potentially critical situations. The results can be used to prevent critical situations by supporting security staff to control the crowd dynamics early enough. In Köfler et al. (2017), a comprehensive system concept was described, which uses both, airborne and terrestrial sensor systems for crowd management tasks in large event situations. In Ibrahim et al. (2016), a general overview of intelligent surveillance systems (ISS) was given as well as concepts and approaches for possible sensor modalities and fusion methods were described. Rai et al. (2019) describes a system approach for an advanced intelligent video surveillance system (AIVSS). The majority of the described developments, systems and research results are focused on specific scenarios with specific technical and operating conditions. A key requirement for security management is the usage of a system, which enables the permanent (24/7) generation of situation information. This requires the synergetic use of multiple sensors-technologies to cover different day / night as well as weather conditions along with an innovative approach in data analysis methods to compensate the advantages and disadvantages of the individual sensors. The key to an optimized operational command support in security management is the ability to get relevant information based on situation specific sensors and data analysis as well the ability to distribute and use this information by all involved stakeholders. The concrete scenarios and the described technology focuses on monitoring for frontier security as well as event security situations. In the national funded KIRAS research project WatchDog (WatchDog, 2020) a concept of a mobile communication and multi-sensor solution (optical, thermal and radar sensors) for security and risk management for outdoor and object security was elaborated, tested and evaluated. This work focuses on the included modules for data gathering, multi-sensor data processing as well as information management for the security companies and involved task forces.

2. System Concept

The goal of WatchDog was the design and prototype development of a multi-sensor system for security scenarios that allow a self-sufficient and, if needed, independent use from commercial communication infrastructure. In order to ensure a 24/7 operational capability, the concept includes, optical, thermal and radar sensors which are synergistically used. Optical sensors (mostly CMOS image sensors) have a very high resolution, can be used with different lenses for different ranges and are generally inexpensive, but the drawback is, that they heavily dependent on lighting and weather conditions. Compared to optical sensory, thermal sensors normally have a very low resolution, recording the heat radiation in the medium to long-wave infrared range emitted by the objects themselves. Therefore, they are not dependent on the lighting situation, but still have a high price tag. Both sensor types are, however, sensitive to weather conditions. For example, water particles in fog not only block the visual spectrum but also reduce the transmission of the atmosphere in the thermal spectral band. To provide extended independence from influences in harsh weather conditions, microwave ground radar modules were modified and integrated for a synergetic usage, knowing that they lack of high resolution. A system employing the complementary devices operating in three different spectral bands (microwave, visual, and thermal infrared) is able to operate in any given

weather condition. It also follows the principle of dual, passive/active intruder detection technique (Dulski et al., 2011). The radar sensor is an active sensor and measures the distance and the relative speed (Doppler-Effect) in a single measuring cycle faster and more precisely than any other sensor principle. As an active sensor, the radar sensor is independent of lighting conditions and only slightly influenced by weather conditions. The angular resolution of a radar sensor is low and object classification is very difficult. An important benefit resulting from using multiple sensors that operate in different spectral bands is the possibility to apply fusion and advanced data processing, which leads to automatic intruder detection and tracking.

In order to provide purposeful support for security forces, in various deployment areas the following modules and functionalities were developed as a proof-of-concept prototype. They were used in tests and demonstrations at various scenarios (shown in Figure 1):

- Multi-sensor masthead combining optical, thermal and radar sensors.

- Energy self-sufficient, transportable carrier system based on a table telescopic mast system, which enables permanent (24/7) use.

- Stable, dynamic on-demand broadband communication solution for secure and robust transmission of multi-sensor data using small unmanned aircraft (SUA) based relay stations, which have been developed as part of the project

- Automation-based data analysis with the integration of optical, thermal, and radar data and based on innovative analysis and sensor fusion methods.

- Modules for decentralized manual or analysis-controlled alignment of the sensors and with the available headlights.

- Operational command support based on high-level visualization and an optimized group and role-oriented data distribution.

- Mobile solutions to support operational command of mobile teams.
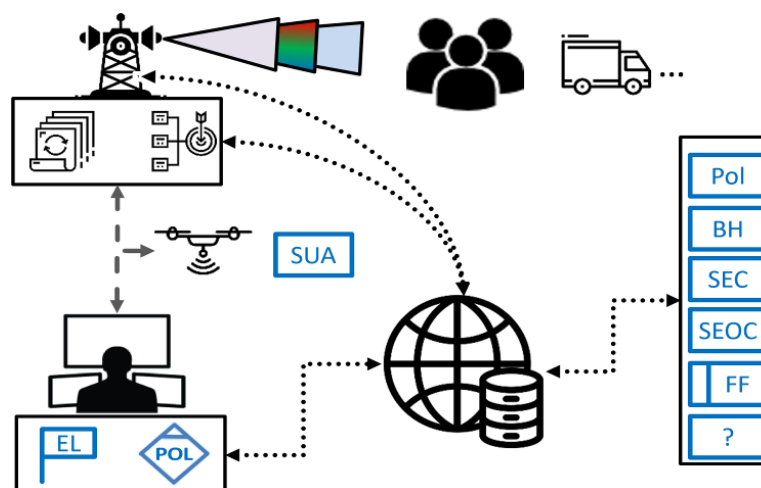
Figure 1: Overview of involved modules and data flow.[4]

Figure 2 gives an overview of the general modules in the experimental WatchDog system approach. The optical, thermal, and radar sensors provide a comprehensive data stream as basis for automatic analysis processes and sensor fusion methods. To integrate the security forces as "human sensor" in

---

[4] SUA - Small Unmanned Aircraft

the information flow, an Android based mobile solution was developed to enable a scenario specific and task oriented management as well as interaction features. The web-based management solution integrates different data and information sources, generates a near real time situational picture (common operational picture) and enables a targeted distribution of information to other organizations, decision makers as well as security forces. Decentralized data access, coordinated using a central data service, allows flexible, and bandwidth aware usage of information products for a large amount of users. Each mast system processes its sensor data locally and actively sends only the small results data and notifications to the central data service. If needed, on-demand access still allows the users to deal with the analysis source data. The networking of the components was achieved using WiFi and LTE technology (Raffelsberger et al., 2019).
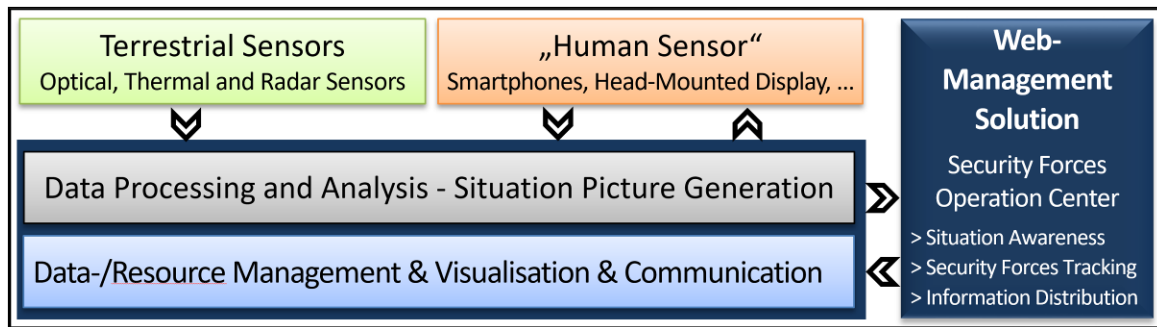
Figure 2: Overview of the system concept.

3. Sensor Platform and System Components

3.1. Self-sufficient multi-sensor data acquisition system

The developments within the project focused on an innovative self-sufficient multi-sensor data acquisition system including optical and thermal cameras, radar sensors and GNSS. As Figure 3 depicts, the sensors data is captured in a synchronized way, so that one data package, consisting of a reading of each sensor, can be used in the following data analysis step of the workflow.
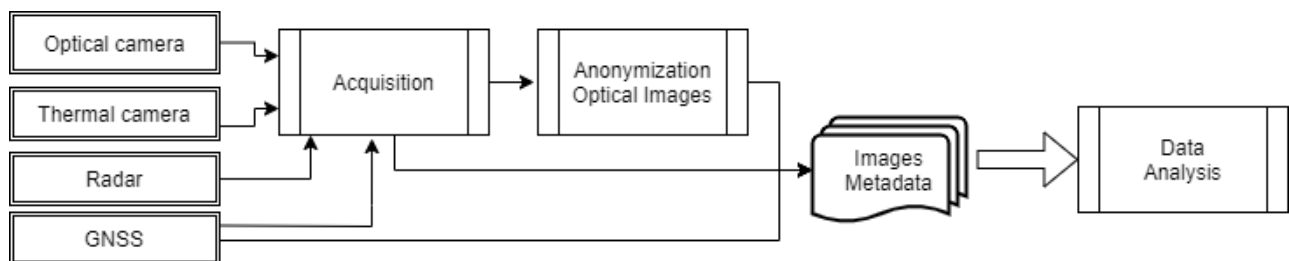
Figure 3: Workflow of multi-sensor data acquisition

The housings and mounting points of each sensors (Figure 4 left) are calibrated mutually so that a single object, which is captured by multiple sensors can be identified as being an object at a certain location. The sensor platform has been designed as an energy self-sufficient module by using mainly infrastructure independent on-site diesel driven power generators, local data processing as well as on-demand and data communication. As optical sensors we used an Axis Q1647 (3072x1728px) and a PTZ Axis Q6128 (3840x2160px) as well as an Optris PI 640 (640x480px) as a thermal sensor. The radar sensor was a custom tailored and enhanced version of the Radarbook by INRAS operating at 24 GHz (INRAS, 2020) and a GNSS.
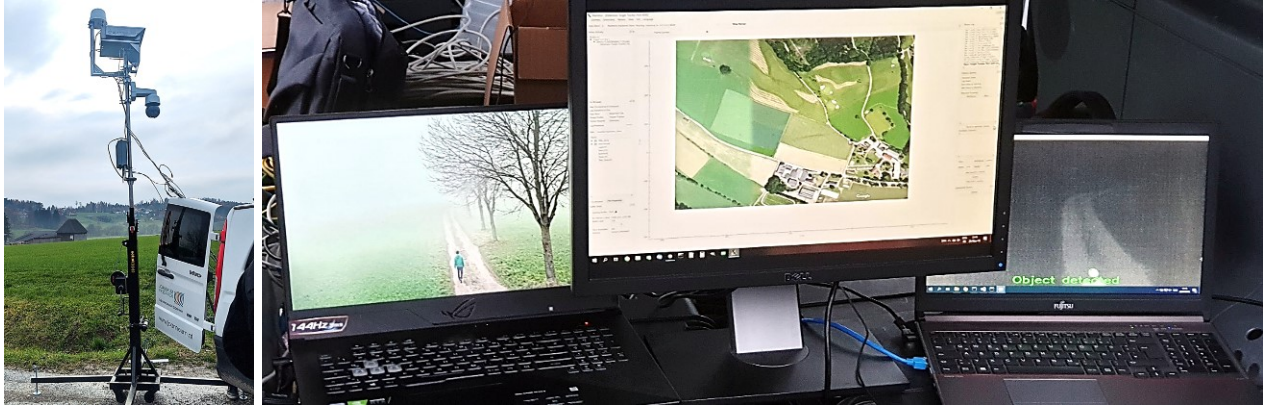
Figure 1: Sensor platform based on the mast system (left) and real time data processing showing optical, radar, and thermal processing directly in the van (right). While for the optical and thermal sensors the input images are shown, for radar the detections are superimposed on Google Maps data.

3.2. Data analysis and information extraction

The data analysis and information extraction developments focused on AI methods and algorithms for optical, thermal and Radar sensor data. Information is extracted separately from each of the modalities and then fused to get stable and reliable measurements. Therefore, a portfolio of algorithms was selected, enhanced, and implemented to serve a variety of scenarios. The main goals were trifold, in particular object detection, object tracking, and object density estimation, the latter also including object counting. For the image-based sensors three object detection paradigms got selected, one based on classical machine learning (Dollar et al., 2010) and two based on deep learning (Redmon and Farhadi, 2018; Ren et al., 2015). While the first is custom tailored for detecting pedestrians, the others can learn to detect any kind of objects. For training, the Common Objects in Context dataset is used (Lin et al., 2018) since it holds a set of useful object categories for security relevant scenarios, like humans, bikes, cars, animals, and additionally object classes like chairs that commonly occur in mass events. For object tracking, the detected objects have to be tracked over time. Thus, the optical flow is extracted over consecutive video frames (Chambolle and Pock, 2010) allowing a re-identification of objects. Density estimation and object counting is again based on two approaches. One based on classical machine learning (Perko et al., 2013) and one based on deep learning (Li et al., 2018). For training two datasets (Zhang et al., 2016; Idrees et al., 2018) are employed allowing best adaption for the project's scenarios. Radar-based detection and tracking build on clustering of the range-doppler maps over time in a process using Kalman filter-based multi-target tracking. This multi-target tracking allows both the differentiation between moving and stationary objects as well as incorrect detections, and the assignment of the individual clusters to these objects. This assignment is an elementary step for the classification (cf. (Wagner et al., 2017)). Finally, the processing results of each data modality is send to a server based on ROS messages building the basis for data fusion.

In order to take advantage of a self-sufficient multi-sensor platform and to obtain stable analysis results regardless of environmental influences, sensor fusion approaches were developed and tested. The single sensor based detections are collected as geometric primitives. In particular, an object detected within an optical or thermal image is represented as a 3D ray originating from the (calibrated) camera's center. Likewise, a radar detection is represented as angle and distance w.r.t. the radar sensor's position. The fusion itself clusters those inputs in 3D and over time. Only if an object is consistent spatially and temporally, it represents a real detection. Thus, the fusion process incorporates multi-modal information to allow stable and robust detections while simultaneously filtering incorrect singular identifications.

3.3. Visualization, Data distribution and operational command support

The use of a multi-sensor platform and the real-time generation of situation-relevant information as the basis for critical decision-making processes, both in the operations center and with the mobile teams on site, requires a management solution geared to the individual user groups in order to be able to efficiently manage data access, processing and information distribution. Two groups of users are being addressed with the user interfaces developed: quasi-stationary operators using the Web GUI and mobile units using a custom Mobile App. As both applications rely on the same data warehouse, it is possible to exchange information and commands between different roles in the network of users. Additionally sensor data streams and analysis results are accessible by these apps.

The HTML5 based single page app hosted by a Node.js server allows role based access to data while being authenticated as a specific user.
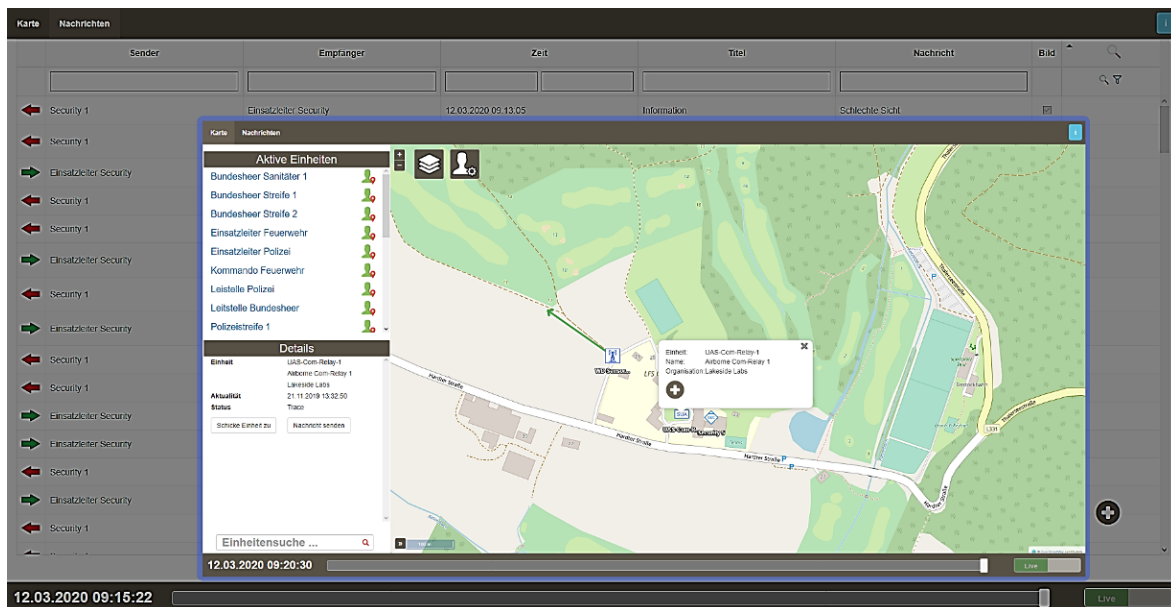
Figure 5: Web App with Message View and Map View (inner floating blue frame)

The surface consists of two primary tabs. On the one hand, this is the Map-View to display geo-referenced information, on the other hand the Message-view for communication purposes (see Figure 5). The Map View presents a selectable base layer like OSM, local tile sources, etc. On top of this, layers represent the current locations of mobile units that automatically send their position and status information using the mobile app. In addition, operating resources like the sensor mast are displayed in their own layers. As shown in the center of the Map View example, the mast is visualized as an icon accompanied by a green arrow, which indicates the current viewing direction of the sensors. The map shows each enabled unit on its geographical position. Units can be found by doing a manual search or using the textual search tool of the unit list-view. For the selected unit, parameters like name, status or age of status information are shown to the user. Based on the object type, the user can interact with the indicated units and resources on map using context menus. This allows sending commands to move to a destination location, sending messages or, in case of the mast, requesting a live stream from the optical camera as well as control the cameras viewport. Another map element is the alert maker represented by a pulsing ring and central symbol. It shows the location of an automatic alert based on automatic alert deduction derived from current sensor data. Next to the described features for real-time operation, the operator can use the time slider to access historic data of the current mission. Using this, temporal and spatial requests are sent to the data warehouse in order to provide the situation for the selected time and allows to analyze past situations just like current data. The second view shows all messages exchanged between operators and mobile units. An operator

can filter and sort messages using the according UI elements. Messages from mobile units, offer the attachment of pictures of situations. The operator can compose new messages as well as answer to one in order to distribute information based on text only and multimedia elements. Each message can include a coordinate for in order to set the information in a spatial context.

Additionally an Android based mobile solution to assist operational command and mobile security forces was developed (Figure 6). It enables security forces to act as "human sensors" and send near real-time information related to their position, activity and status to help decision makers gather a comprehensive common situational picture. Visualization techniques on head mounted displays as well as their practical usage were evaluated. Based on this, the mobile solution was adapted to directly run on a head mounted display, supporting navigation, displaying instructions in the field-of-view and allowing security forces unobtrusive, hands-free operation in the field. The application implements an efficient communication paradigm enabling fast and intuitive data and instruction exchange between operational command and mobile units. It incorporates automatic messaging, broadcasts or group-based messaging, and an adaptive role-based user interface (UI) with role based and situation related auto-filled contents. The application design and UI development considered requirements and common usage patterns of security forces like procedures from different digital radio devices as well as easy and role based information access. The mobile operational picture module implements a map component displaying other mobile forces, relevant infrastructure and multi-sensor mast positions. Next to that, it allowed the interaction with the named entities like direct access and remote control of cameras and lighting devices. Features like quick in-situ response to geo-oriented information in dynamic situations is also one of the evaluation topics as it directly effects operational tactics.

The goal was to create an easy and highly usable application, which supports and facilitate the practical work of security personal as well as first responders. In addition, it was designed to extend current operation procedures and operation tactics by optionally providing the described features and functionalities.
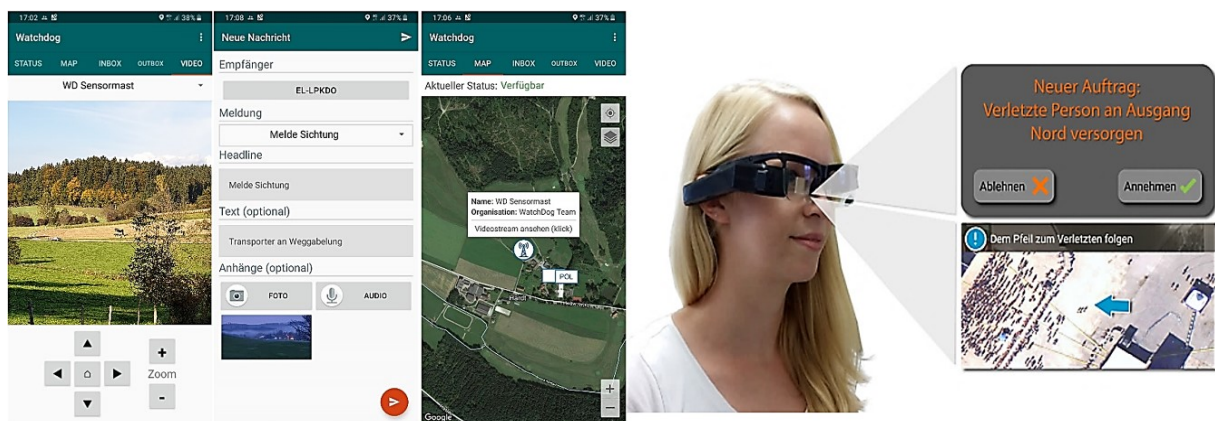
Figure 6: Mobile information system (left) and "head mounted display" solution for flexible use with mobile teams (right).

4. System Tests and Demonstrations

Multi-modal data was acquired on multiple events, including several test scenarios but also on real events, like the Frequency music festival in Austria.

4.1. Data acquisition and single data analysis

AI-based object detection, tracking and density estimation are demonstrated on each single data modality with the objective of extracting the best possible result for each modality. For example, the detector from Redmon and Farhadi (2018) produces better detections on optical data, while the detector from Ren et al. (2015) performs better on thermal imagery. Exemplary object and motion detections are shown in Figure 7 for one of the test scenarios. The figure also depicts radar-based object tracking showing multiple extracted trajectories of pedestrians and cyclists.

Figure 7: People detection from optical (left) and thermal data (middle). Radar-based detection trajectories (right) with image courtesy of INRAS and Johannes Kepler Universität Linz.

4.2. Data acquisition and information extraction at Frequency festival 2019

At Frequency festival 2019 all three data modalities were jointly exploited. Figure 8 and 9 shows exemplary density estimation and counting, where the human density is color coded (blue - low; green - middle; red - high density) and superimposed to the original optical image. The figures also depicts thermal detections based on motion estimation and object detection and tracking in optical video data.
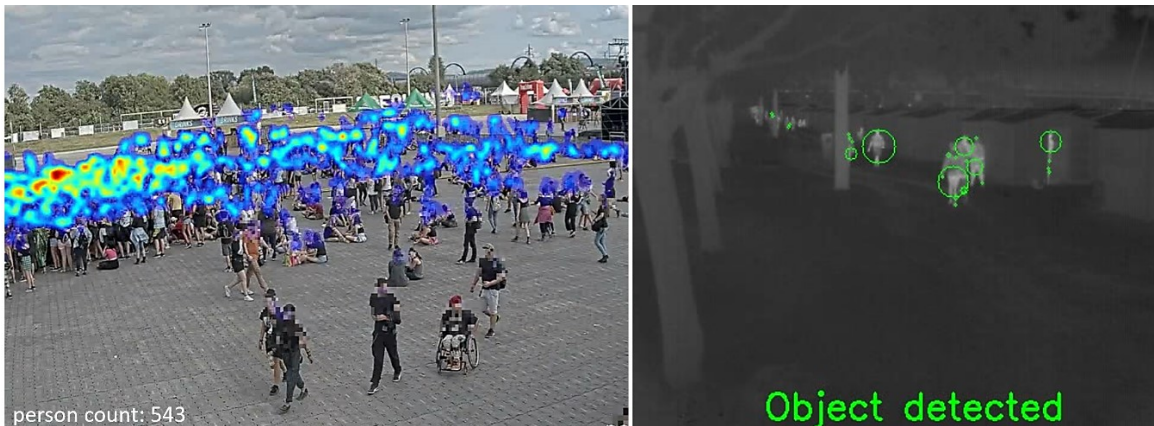
Figure 8: Frequency 2019 – automatic crowd density estimation and counting of humans based on optical video data (left), as well as, person detection in hut storage based on thermal data (right).

Overall, thermal images yield stable detection, especially in the night, and are also useful to detect small changes. The optical image-based object detector aims to identify and track multiple object categories and is very stable at daytime, however objects need to be larger for robust detection. The radar sensor (not shown) always yields good detections however at a low angular precision. Thus, the data fusion step allows to emphasize the benefits of each data modality.

Figure 9: Frequency 2019 - automatic person detection based on optical data (bottom). Individual detections before sensor fusion are shown.

5. Conclusion and Outlook

The recent past has shown that dynamically changing characteristics of security and emergency management requirements have a strong impact on the work and resources of security forces to ensure the security in public spaces and critical areas. The concepts and prototypical system solutions described in this paper enable directed assistance in the deployment of security forces in the context of frontier protection scenarios, security management at large events and critical infrastructure protection. In order to ensure a 24/7 operational capability the WatchDog system integrates complementary sensors operating in different spectral bands (microwave, visual and thermal infrared). AI-based methods for multi-sensor data analysis enable the generation of a near real-time situation picture. Web-based and mobile management solutions support the integration and fusion of different data sources and the targeted distribution of information to support time critical decision processes.

In general, this paper shows the potential and benefits of an innovative multi-sensor solution. Although several promising results have been obtained, further research activities are needed to optimize the sensor performance for a greater system robustness and fully utilize the potential of the different, complementary sensors based on the development of AI-methods for a multi-sensor data analysis and fusion. With decreasing prices for specific sensors and processing devices, it will be a very important aim for future research activities to consider investigating and developing an intelligent, efficient low cost multi-sensor surveillance system.

6. Acknowledgments

7. References

Almer, A., Perko, R., Schrom-Feiertag, H., Schnabel, T., & Paletta, L. (2016). Critical Situation Monitoring at Large Scale Events from Airborne Video Based Crowd Dynamics Analysis. In Geospatial Data in a Changing World pages 351–368. Springer, AGILE, [Best Paper Award].

Chambolle, A., & Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision, 40(1), pages 120-145.

Dollár, P., Belongie, S., & Perona, P. (2010). The fastest pedestrian detector in the west. In: Electronic Proceedings of the British Machine Vision Conference (BMVC), No. 68.

R. Dulski, M. Kastek, P. Trzaskawka, T. Piątkowski, M. Szustakowski, & M. Życzkowski (2011): "Concept of data processing in multi-sensor system for perimeter Protection", Proceedings of SPIE - The International Society for Optical Engineering, DOI: 10.1117/12.883965.

Essendorfer B., Monari E. & Wanning H. (2009): "An Integrated System for Border Surveillance", International Conference on Systems, DOI: 10.1109/ICONS.2009.27.

Idrees, H., Tayyab, M., Athrey, K., Zhang, D., Al-Maadeed, S., Rajpoot, N., & Shah, M. (2018). Composition loss for counting, density map estimation and localization in dense crowds. arXiv:1808.01050.

INRAS (2020): Radarbook; http://www.inras.at/en/products/radarbook.html (accessed May 2020).

Köfler A., Pammer-Schindler V., Almer A., & Schnabel T., (2017): "Supporting decision making in security forces' command centre at large-scale events via video-based situation monitoring and base data supply". i-KNOW.

Li, Y., Zhang, X., & Chen, D. (2018). CSRNet: Dilated convolutional neural networks for understanding the highly congested scenes. In IEEE Computer Vision and Pattern Recognition, pages 1091-1100.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In European Conference on Computer Vision, pp. 740-755, Springer, Cham.

Mritunjay Rai, Agha Asim Husain, Tanmoy Maity, & Ravindra Kumar Yadav (2018): "Advance Intelligent Video Surveillance System (AIVSS): A Future Aspect". IntechOpen, DOI: 10.5772/intechopen.76444.

Perko, R., Schnabel, T., Fritz, G., Almer, A., & Paletta, L. (2013). Counting people from above: Airborne video based crowd analysis. In Workshop of the Austrian Association for Pattern Recognition, number 37, pages 1-9, arXiv:1304.6213.

Raffelsberger, R., Muzaffar, R. & Bettstetter, C. (2019) A Performance Evaluation Tool for Drone Communications in 4G Cellular Networks. International Symposium on Wireless Communication Systems (ISWCS), IEEE, pages 218-221. doi: 10.1109/ISWCS.2019.8877360

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pages 91-99.

Sutrisno Ibrahim (2016): "A comprehensive review on intelligent surveillance systems", Communications in Science and Technology 1, pages 7–14.

WatchDog (2020): Mobile communication and multi-sensor solution for security and risk management outdoors and in object protection; https://projekte.ffg.at/projekt/1698680 (accessed April 2020).

Wagner, T., Feger, R., & Stelzer, A. (2017). Radar signal processing for jointly estimating tracks and micro-Doppler signatures. IEEE Access, 5, pages 1220-1238.

Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-image crowd counting via multi-column convolutional neural network. In IEEE Computer Vision and Pattern Recognition, pages 589-597.

THE RELEVANCE OF TAILORED TERMINOLOGIES FOR LARGE SCALE INTERNATIONAL PROJECTS

Esther Kähler

DIN
Esther.Kaehler@din.de

Georg Neubauer

AIT Austrian Institute of Technology GmbH
georg.neubauer@ait.ac.at

Marcel van Berlo

TNO The Netherlands Organisation for Applied Scientific Research
marcel.vanberlo@tno.nl

Tim H. Stelkens-Kobsch

German Aerospace Center DLR
tim.stelkens-kobsch@dlr.de

Keywords

terminology, semantics, taxonomy, vocabulary, ontology

Abstract

Within DRIVER+ a terminology was developed including the most relevant terms and definitions applied within the project. This terminology was dedicated to enhancing shared understanding within the project team as well as to support an improved understanding of the output of the project in the broader crisis and disaster management community. This paper describes how the project terminology was built based on existing terminologies and the added value of such an approach.

1. Introduction

In the frame of international research projects, a multitude of actors with different cultural, organizational, national and educational background come together in order to achieve one or more targets. Under such conditions it is essential to ensure a common understanding of the actors, as there is otherwise a high risk of failure in achieving the common goals. Additionally, the naming and defining of newly developed products of the project is needed in order to build a shared understanding

of these outcomes. In this light, the EU-funded project DRIVER+ decided to establish a project terminology of key terms and associated definitions in order to enhance a common understanding within the project team and to contribute to a shared understanding within the broader crisis and disaster management community within Europe (DRIVER+ (2020). This procedure is only known from a few projects and in this area, the DRIVER+ terminology is the most extensive and result based known by the authors.

2. Method

2.1. Example for Building a Terminology

Within DRIVER+ an innovative and well-structured process was followed to build the DRIVER+ terminology. A Terminology Working Group (TWG) was established dedicated to identifying central terms applied in the project and to provide suitable definitions for those terms. A TWG should consist of people having a good overview of the project and having access to relevant terminologies and standards. In DRIVER+ the group consisted of the scientific coordinator, the quality manager, and further partners having expertise in semantics and standardisation. Depending on the topic, different participants of the project supported the activities of the TWG.

As a starting point for a project terminology an existing terminology, such as UNISDR (2015) or ISO 22300 (2018) can be used. From there on the most relevant terms for the project and their definitions can be identified. The first step to establish a basis version of the project terminology is to identify the relevant terms of this initially available list.

Next, a process of continuous update of the terminology has to be set up: In the cases when an update for specific terms was requested in DRIVER+, this resulted from an input of a project partner or a member of the TWG. The TWG forwarded the request to the Project Coordination Team (PCT), consisting of the project coordinator, the technical coordinator, the external cooperation manager and the quality manager and in case of acceptance of the request the TWG took the following actions:

1. The TWG examined if an adequate definition could be found in UNISDR (2015), ISO 22300 (2018) and the ISO database (2020).

2. If one or more adequate definitions were found in these sources, all definitions were proposed to the PCT together with a recommendation for a selection.

3. If no adequate definition was found, other already available sources were examined without giving any ranking to the PCT in case of identification of adequate definitions. These sources were:

   a. SDSIE (2017): Service de défense, de sécurité et d'intelligence économique, Crisis & disaster management terminology translation.

   b. IFRC (1991): Code of conduct – International Federation of Red Cross and Red Crescent Societies.

   c. ICRC (2013): Professional standards for Protection Work.

   d. HM Government (2013): Emergency Response and Recovery Non-statutory guidance accompanying the Civil Contingencies Act 2004.

In case all examined sources did not provide any or adequate definitions, additional sources could be examined, or self-defined definitions could be taken. The whole process is illustrated in Figure 1.
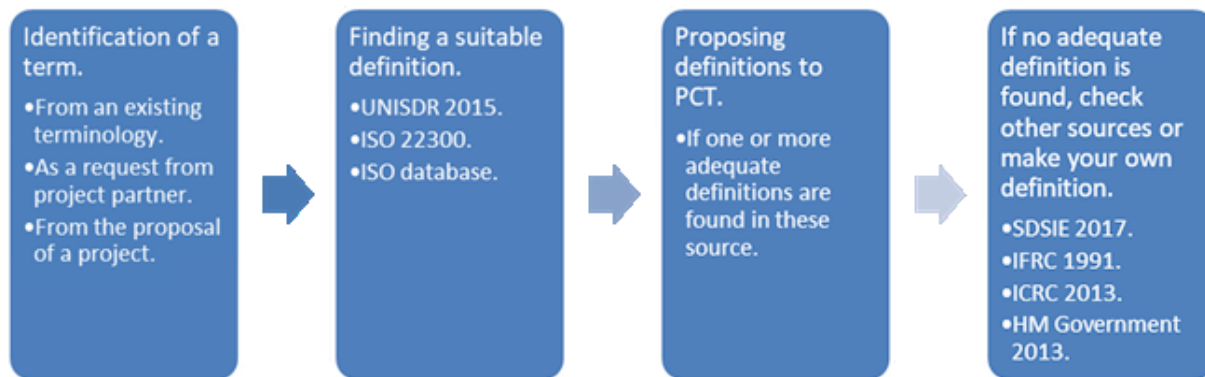
Figure 1: The search process for definitions according to the DRIVER+ methodology

In general, recognized and well accepted definitions were given priority compared to self-defined DRIVER+ definitions. However, although sometimes available definitions were suited for the project purposes, an additional comment to the definition was deemed helpful in order to make the DRIVER+ context more specific. This was done by adding a so-called DRIVER+ note in addition to the definition.

On this basis the DRIVER+ terminology was continuously updated, published on the DRIVER+ public website and included in the Portfolio of Solutions (PoS) which was also developed within the DRIVER+ project. The PoS is a repository of solutions for crisis and disaster management allowing to match the features of the solutions with the crisis management gaps defined by practitioners (PoS, 2020). Whenever a term of the DRIVER+ terminology is used in the description of a solution, the PoS provides the corresponding definition from the DRIVER+ terminology via mouse over function. The DRIVER+ terminology aims to contribute to a shared understanding in Crisis Management and to be sustainable simultaneously. Therefore, it is also translated in eight languages, i.e. German French, Dutch, Italian, Spanish, Polish, Estonian and Swedish. The translations were done to further increase the mutual understanding. The chosen languages cover a large part of the European Community and help the DRIVER+ Centres of Expertise (CoE) in their activities to share experiences with regards to supporting the capability development and innovation management of European practitioner organisations (CoE, 2020).

2.2. Acronyms and Abbreviations

In addition to the project terminology a list of acronyms and abbreviations used in DRIVER+ was established. The catalogue of acronyms and abbreviations was set up in order to foster achievement of a common understanding on names for systems, tools, institutions etc. in the project. Like a terminology a mutually agreed register of abbreviations and acronyms helps to achieve seamless communication in a project and to avoid possible misunderstandings on specific topics. The different lists of acronyms and abbreviations in each deliverable were used as input for the global list available to the project and beyond. After reception of the first deliverables the Terminology Working Group began to compile the list and provided it to the project consortium. Thereafter every new occurrence of acronyms and abbreviations was monitored and registered in order to avoid duplications.

The updated terminology (terms, definitions) and the list of acronyms and abbreviations are published on the public project website.

3. Results

The DRIVER+ terminology consists of 147 published terms and 556 acronyms and abbreviations (Terminology, 2020). 17 of 147 definitions are extended by a DRIVER+ note. Most definitions were found in the standardisation documents via the ISO online browsing platform (ISO database, 2020): 16 definitions come from ISO 22300:2018 Security and Resilience – Vocabulary, and 64 definitions from other standards. UNISDR was the source for 9 definitions. 34 definitions are initial DRIVER+ definitions, developed by the project and 24 terms were defined by different sources in the area of CM. The results are shown in Figure 2.
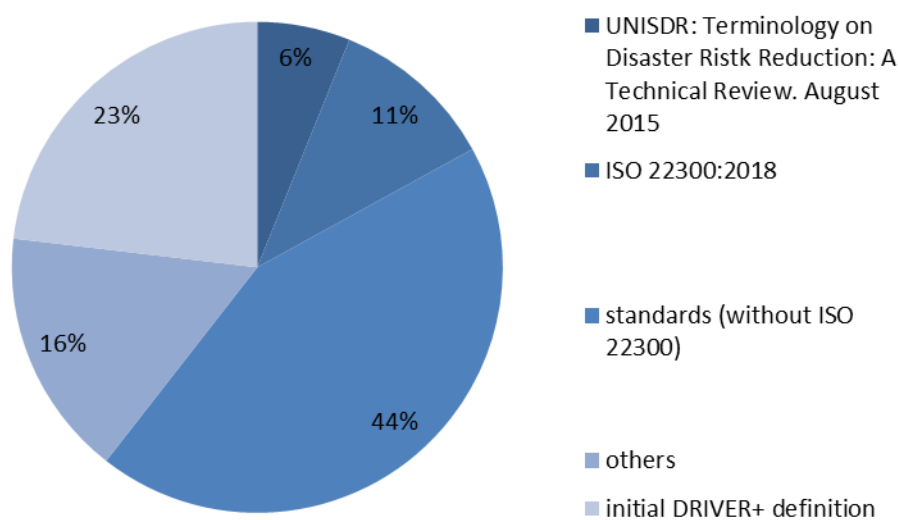
Figure 2: Overview of the sources used for the DRIVER+ terminology

The main part of the terminology was translated in eight different European languages in order to enhance the outreach in the European crisis and disaster management community. Due to the fact that a few terms were added to the terminology after the execution of the translations, translations are not available for all of the terms.

The Crisis Management Innovation Network Europe (CMINE) has established a group on terminology (CMINE, 2020). Here the first discussions started about definition of existing and new terms utilised in the Crisis Management domain.

4. Discussion and Conclusions

The terminology of DRIVER+ turned out to be a very helpful tool for the execution of the work within the project. However, our intentions are not limited to DRIVER+ as the authors intend to contribute to an improved common understanding of the international crisis and disaster management domain. The outcomes of the project are freely available for any organisation, standardisation or other initiatives. An example for this is the European H2020 project STAMINA on pandemic management that will apply the DRIVER+ terminology. Another example of an interaction is related to the development of the standard EN 17173 (2019). Eight of the initially defined terms by DRIVER+ were submitted to and discussed with the national standardisation committees in Germany and the Netherlands to be involved in the EN 17173 European CBRNE glossary by the responsible European standardisation organisation in 2019. However, this initiative was not successful to the very

end, as the terms were recognised as not fully applicable for the CBRNE based glossary. Another example of an interaction is related to the Interactive Terminology for Europe (IATE, 2020). The IATE platform has decided to include selected terms of the DRIVER+ terminology to expand it on the area of disaster management. At the time of publishing this paper the process is still ongoing to select the most suitable terms and definitions in order to avoid duplications with current IATE terms as much as possible.

The question arises if a terminology shall be recommended for any project or limited to large scale international projects involving experts from several organisations and domains. The authors recommend using a specific terminology in every project to enhance a shared understanding, but in small, national projects limited to one domain, it might be sufficient to select an existing terminology such as ISO 22300 and apply it for the work of the project. The larger and more complex a project is getting, the more justifiable is the invested effort to set up a project specific terminology.

A perspective of a follow up activity is the establishment of a digital thesaurus. The international and national disaster management experience clearly demonstrate that ambiguities and fuzziness in communication at both operational and staff level are one of to the major challenges of crisis and disaster management. The main reason is the lack of mutual understanding of key terms among the different actors. There are multiple reasons for this situation, encompassing different histories and mandates of organisations like fire brigades and police, as well as different legal frameworks. The complexity increases even further in case communication takes place across borders. Within a project, an organisation or in the frame of another type of initiative the actors can agree on one single terminology. But in particular in cross border scenarios the agreement on one terminology cannot be expected. To overcome these challenges the DRIVER+ partners AIT and DIN together with multiple stakeholders such as the Austrian Ministry of Interior and Austrian Standards are currently preparing a concept for a digital thesaurus for crisis and disaster management. This concept includes the possibility to compare terms and definitions from different vocabularies and also encompasses the option for the analysis of semantic similarities of definitions. As a basis a system-internal database with the most important national and international terminologies is planned to be made available to national and international stakeholders. The DRIVER+ terminology has the potential to become the reference terminology of such a thesaurus.

5. Acknowledgment

6. References

CMINE (2020), Crisis Management Innovation Network Europe, https://www.driver-project.eu/cmine-project/, retrieved 13.06.2020

CoE, DRIVER+ Centre of Expertise, https://www.driver-project.eu/centres-of-expertise-coe/, retrieved 13.06.2020

DRIVER+, https://www.driver-project.eu, retrieved 13.06.2020

EN 17173 (2019), European CBRNE glossary

HM Government (2013), Emergency Response and Recovery, Non statutory guidance accompanying the Civil Contingencies Act 2004, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253488/Emergency_Response_ and_Recovery_5th_edition_October_2013.pdf, retrieved 13.06.2020

IATE European Union Terminology, https://iate.europa.eu/search/standard, retrieved 13.06.2020

ICRC (2013), Professional standards for Protection Work, https://www.icrc.org/eng/assets/files/other/icrc_002_0999.pdf, retrieved 13.06.2020

IFRC (1991), The Code of Conduct for the International Red Cross and Red Crescent Movement and Non-Governmental Organisations (NGOs) in Disaster Relief, 1991, Prepared jointly by the International Federation of Red Cross and Red Crescent Societies and the ICRC, http://www.ifrc.org/Global/Publications/disasters/code-of-conduct/code-english.pdf, retrieved 13.06.2020

ISO 22300 (2018), Security and resilience – Vocabulary

ISO database, https://www.iso.org/obp/ui/#search, retrieved 13.06.2020

PoS, DRIVER+ Portfolio of Solutions, http://pos.driver-project.eu, retrieved 13.06.2020

SDSIE – Service de defense, de securite et d'intelligence economique, Crisis & crisis management, terminology, March 2017

Terminology (2020), DRIVER+ Terminology, https://www.driver-project.eu/driver-project/terminology/, retrieved 13.06.2020

UNISDR (2015), Proposed Updated Terminology on Disaster Risk Reduction, August 2015, https://www.preventionweb.net/files/45462_backgoundpaperonterminologyaugust20.pdf, retrieved 13.06.2020

DIGITALIZATION, SUSTAINABLE DEVELOPMENT, CORONAVIRUS AND CONSUMER BEHAVIOR CHANGE

Lenka Švecová, Galina Ostapenko, Jaromír Veber

Masaryk Institute of Advanced Studies Czech Technical University in Prague
lenka.svecova@cvut.cz, galina.ostapenko@cvut.cz, jaromir.veber@cvut.cz

Keywords

Abstract

The purpose of this study was to examine the global trends that have emerged over the last twenty years, such as digitization reflected in e-commerce, sustainable production and consumption, and the latest global challenge - coronavirus pandemic with the intention of understanding their impact on consumers' behavior change. The study employs the qualitative research using content analysis, the secondary analysis of existing statistical data and the primary research, which was conducted among the younger generation in Prague, Czech Republic in April, 2020. The study confirmed that e-commerce-based sales models and sustainable production and consumption play an indispensable role, having a sensitive influence on people consumer purchasing behavior. Certain shifts in changing consumer behavior can be expected as a result of the coronavirus pandemic. The findings are beneficial for manufacturers and distributors of products and services.

1. Introduction

Digitalization and sustainable development are two of the most powerful market influences in today's business landscape, changing consumer behavior. With the advancement of digitalization, new distribution models are emerging, which are tied to online platforms. They provide the choice and comparison of products in terms of quality, price and delivery conditions. Digitalization change retail settings (Ritzer & Nathan, 2010) and affect the level of customer satisfaction (Poncin & Mimoun, 2014) Today the future depends on e-commerce, new way of doing market research and using new digital strategies. Extreme development of Internet purchasing is associated with the younger generation growth, known as Millennials (born between 1981 and 1996) and "post Millennials" - Gen Z (born after 1997). Generation Y is the first global generation connected by the Internet and social media (Espinoza & Ukleja 2016). Gen Z is more pragmatic and realistic generation of consumers expects to access and evaluate a broad range of information before purchases (Francis & Hoefel, 2018). Determining their preferences and the abilities to become sustainable lifetime consumers, as well as consumer behaviors (especially, purchasing attitudes and patterns) changes under the influence of global trends and challengers has become an important focus of consumer research.

Partly as a result of consumption patterns, society and businesses are confronted today with sustainability factors, including environmental degradation, carbon emission, climate change. In connection with enhanced environmental awareness, higher demands are placed on sustainable production and consumption (SPC) towards zero waste and zero carbon-footprint (New Green Deal),

implementing the circular economy (Svecova, Ostapenko, Veber, 2019). Consumer behavior is key to the impact that society has on the environment. Many researchers found that consumers value sustainable attributes (Trudel & Cotte, 2009), consumers have awareness, concern and positive inclination towards green products (Mozer, 2015). However, the awareness of people and practical applications of new concept of zero waste circular economy are still low in business and among the younger generation (Svecova, et.al., 2019). Therefore, it is important to monitor the dynamic in consumer behavior changes towards to sustainable consumption. Changes in consumer behavior, especially in developed countries reflected how society and market environment are changing. The customers are becoming more powerful in making their own purchasing decisions. The reason for shopping is no longer just the necessary needs, but because it is fun, the goods are cheap enough, they can be different from others in the short term, therefore in practice of shopping and buying it is often observed the impulse purchasing behavior. Impulse purchases represent almost 40% of all the money spent on e-commerce. More than 52% of millennials were more likely to make impulse purchases than any other generation (Amos, et al., 2014). Many brands try to reach consumers on this level through emotional marketing, partly in response to the research works by Nobel prize psychologist Daniel Kahneman, who showed that most decision-making is "irrational". However, in some cases, these emotional marketing efforts have failed. An example is the stress situation, or economic crisis and recession. The individuals are more strategic about how they spend their money under stress, spending more on necessities and less on discretionary items.

The widespread outbreak of coronavirus, which causes the disease COVID-19, is a global pandemic, and it is a real stress situation for businesses and customers. As coronavirus continues to dominate daily lives and newsfeeds, consumers, also businesses are understandably anxious – not just about the disruption to their day-to-day, but about the longer-term consequences. Consumers are shifting their spending habits, they are becoming more reserved about large and expensive purchases and spending money; due to necessity and convenience they are shopping mostly online. Some product categories and companies, mostly in e-commerce business are actually growing. In Accenture's article 'Navigating the impact of COVID-19 in Consumer Goods' it was stated that this ongoing pandemic will permanently alter consumer behaviors that will result in lasting structural changes to the consumer goods industry (Accenture,2020).

2. Methodology

The authors believe that new trends that have emerged over the last twenty years, such as digitization, sustainable development, as well as changes in the lifestyle of members of society, have an impact on both customer purchasing behaviour and decision-making process of producers and suppliers, including their distribution. The authors focus on examining the hypotheses that e-commerce-based sales models and sustainability in the production and product consumption (lower energy requirements, lifetime extension, the circular and green economy principles) have had an impact on changes in shopping behaviour in recent years. The authors are also interested in purchasing preferences with regard to the environmental impacts of product production (cultivation), the origin of products, the customer attitudes to local brands. The methodological tools for examining the research questions and hypotheses was the qualitative research using content analysis and secondary analysis of existing statistical data. In addition to qualitative search, the primary research of consumer purchasing behavior with emphasis on contemporary trends related to digitization and sustainable development, including the questions aimed to discover the short term and long term impact of coronavirus – COVID - 19 on consumer behavior change was conducted in April 2020. The target group was the younger generation (Millennials and Gen Z). These generations are the most tech-savvy generations; the majority of them are smartphone users with digital applications; at least this is a large demographic group that will determines purchasing behavior over the next few years.

According to a typology of different target marketing groups based on Sinus –Milieus® 2018 model, among them also: Cosmopolitan Avant-gardes, interested in pushing boundaries and changing traditional systems, increasingly relying on the internet to network globally, and the Adaptive Navigators, who are happy to adapt to changing requirements. Both groups are the quickest to grow; their approaches to current challengers are indicative of future trends (Sinus-Milieus, 2018).

The Questionnaire with 28 questions was constructed to find out attitudes, preferences and opinions of young people to contemporary trends and challengers facing them. We focused on investigation the young people attitudes and their preferences to on-line shopping; some questions were asked about **brand** purchase behavior; others were asked to find out the young people attitudes towards the sustainable development and environmental issues. Concerning the coronavirus impact, the questions were proposed with the aim to know how the behavior of young people has changed since the beginning of quarantine. The structured questionnaire was formalized by Google Forms and was shared online, so participants could access it. Together 130 respondents were approached, 98 were returned (75,3%). The results from secondary and primary researches are presented below.

3. Results

3.1. Analyzing Data

E-commerce undoubtedly represents the current model of distribution, which uses the Internet, information platforms, allows mobile communication, etc. Online shopping is one of the most popular online activities worldwide with global e-retail sales reaching 3.5 trillion U.S. dollars in 2019. In 2020, the number of digital buyers is expected to be 2.05 billion, that makes 26.28% of the 7.8 billion people in the world (Statista, 2020).The customers use on-line shopping because of money saving, convenience, speed delivery and the possibility of product comparison.

Table 1 Development of e-shops in the Czech Republic (CR) and Germany (BRD)

Year	B2C-ecommerce revenue in billion euros (CR)	Year-on-year growth percentage (CR)	B2C-ecommerce revenue in billion euros (BDR)	Year-on-year growth percentage (BDR)
2014	2,6	15,5	35,6	10,0
2015	3,2	18,0	39,9	10,7
2016	3,8	15,7	44,2	9,7
2017	4,5	15,5	48,9	9,6
2018	5,3	15,1	53,8	8,7
2019	6,3	15,8	57,8	7,2

The above table 1 shows, that the Czech Republic is the fastest-growing e-commerce market in emerging Europe. The country has one of the highest numbers of online stores per capita in Europe. The B2C-e - commerce in the Czech Republic increased from 2 billion euros in 2014 to 6,3 billion euros in 2019 with more, that 15% annual growth rate (Statista, 2020). The strongest segment of customers shopping through e-shops are young people aged 25-34 years, then the segment in the age category 16-24 years, followed by the age group 35-44 years, and at least followed by seniors. Online sales are growing much faster than brick-and-mortar sales, causing a steep drop in business of physical stores. About 60% of respondents in the Czech Republic expect a brick and mortar store to have an e-shop. A way to reduce transaction costs for starting e-shops is offered by an intermediary e-shop that does not offer its own goods, but an online communication space for various small shops.

With the coronavirus quarantine period, some contradictory tendencies can be observed in the Czech Republic e-shops. Up to 40 percent of Czech online stores are affected by supply outages from China. Due to this fact, some online stores have testified up to a third decrease in turnover from the beginning of coronavirus quarantine. Domestic online stores, which sell goods coming directly from China experienced supply disruptions in 70% of cases. During the quarantine period, more elderly than younger generations started shopping online: young people increased spending by 34%, families by 47%, and the elderly by 52%. On the other hand, the unusual e-shops boom is observed in March and the beginning of April 2020: consumer purchases increased by 44% compared to the same period of the last year; food product purchases increased by 215%; pharmacy purchases increased by 286%. The preference for payment cards following by significant (by one quarter) fall in cash withdrawals from ATMs has also been demonstrated (Slovenská Spořitelna CZ, 2020).

Exploring the sustainable development issues we focused on current positive practical signals in sustainable production and consumption, especially agriculture and food. These natural products are intended to provoke a greater customers' interest. Organic food in the EU is the result of organic farming, which relies on techniques such as: crop rotation, green manure, compost, and biological pest control; it uses fertilizers and pesticides, if they are considered natural, but it excludes the use of various methods (synthetic petrochemical fertilizers and pesticides, plant growth regulators, antibiotics, genetically modified organisms and nanomaterials).

Table 2 Share of total organic area from total utilized agricultural area in EU (in 2018)

Country	Share	Country	Share	Country	Share
Austria	24,1	Czech Republic	14,8	France	7,0
Estonia	20,6	Slovakia	9,9	Hungary	3,9
Sweden	20,3	Spain	9,3	Poland	3,3
Italy	15,5	Germany	7,3	Bulgaria	2,6

Table 2 above shows that the Czech Republic is among the top of European Union countries with the highest proportion of land under organic agriculture from the total area of agricultural land. Within the framework of the EU, the CR takes 5th place after Austria, Sweden and Estonia, with 14,8% of the area being organically farmed (Statista,2020). Environmental benefits of organic production have clearly contributed to building a positive consumer attitude towards organic. According to the several surveys and statistical information from "Think tank EP", "Eurostat", "Eurobarometer", European consumers appreciate the production with very limited use of pesticides, fertilizers and antibiotics (79%), environmentally friendly production (78%), higher standards for animal care (76%), higher quality and safety (70%). Organic products can be recognized by their labels, with words organic, eco or bio, the EU organic logo (Euro Leaf) and the certification authority number. We can find a number of other labels/certificates: sustainable agriculture certificates (offered by UTZ and RAinforest Alliance); RSPO designation for palm oil produced from organic plantations; Marine Stewardship Council (MSC) designation for sustainable fishery products; SAN / GlOBALG.A.P certification (unified in the future) focused on safe food, FAIRTRADE, labeling of six products (bananas, cocoa, coffee, cotton, sugar and tea); OEKO-TEX certification label for textiles, made only from non-hazardous materials, others. However, the awareness of these initiatives is still embryonic for the general public as well as for the professional community, while retail chains are more aware of the fact that the food they sell has signs of sustainable agriculture.

The results of secondary research confirmed that changes in socio-economic development coming with a new millennium have several important impacts on consumer behavior, among which are:

- The advent of digitization and related e-commerce is reflected in the growth of sales through e-shops. The Czech Republic e-shop revenues increased by more than three times during the last 5 years, showing the biggest in Europe average annual growth (more than 15 %). The contradictory tendencies in e-shops in the Czech Republic during coronavirus period have been demonstrated;

Sustainable development is reflected in sustainable production and consumption. Different approaches have a common denominator and act as a competitive advantage for producers and consumers. At present, it aims to ensure the positive environmental benefits of sustainable production. In order to increase sustainable production and consumption it should be in presentation of marked goods, promotion of organic food consumption in its marketing campaigns, and should be inspected by consumers.

3.2. The Results of primary research on consumer behavior of young people

The primary research of consumer purchasing behavior with emphasis on contemporary trends related to digitalization and sustainable development and partly also influenced by coronavirus restrictions was conducted in mid-April 2020. The majority of respondents were young people from Prague, Czech Republic. 87% of our 98 participants were between the ages of 18 and 26, among whom 58% were female and 42% were male. The most important results of this survey can be summarized as follows:

More than 46.3% of respondents renew their mobile phone within three years, 18.9% within two years, 28.4% within four to five years;

More than 17% of respondents purchased in e-shops weekly, about half of respondents - per month in normal regime, before introducing coronavirus quarantine. Prospectively, 76.8 % of respondents expect to use e-shops for purchases even more in the near future;

To the question: "What is a big inconvenience to you in a process of online shopping?", - 59% of respondents answered that this is the fact that goods cannot be physically viewed and assessed; for 27,4 % of respondents this is the cost of transport; for 13.7 % of respondents the big inconvenience is late or missed delivery;

3,2 percent of respondents use credit cards to pay for almost all purchases; 23,2 % use them for payment of three - quarters of their purchases and 20 % use credit cards paying the half of all their purchases (Figure 1).

Question: How can you estimate the proportion of your credit-card payment in all purchases?
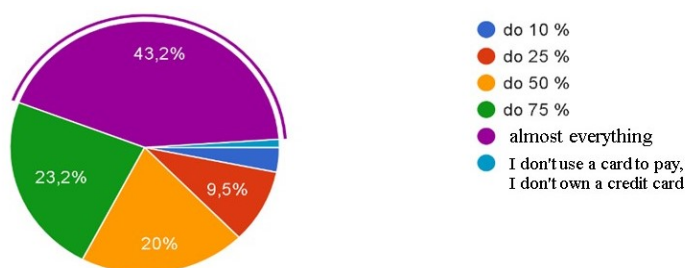
Figure 1. Overview of respondents, using credit-card in their purchasing

During coronavirus quarantine period about 60% of respondents reported that they have used online shopping to purchase clothing, food and electrical goods (50%, 41,5% and 35% of respondents respectively), followed by drugstores and goods for home and garden (Figure 2). To purchase food

online they used online stores with delivery, 29% used rohlik.cz and 19% used kosik.cz. The Czech brand "Klasa" was recognized and appreciated by all respondents.

Question: I used the e-shop during the coronavirus period to purchase:
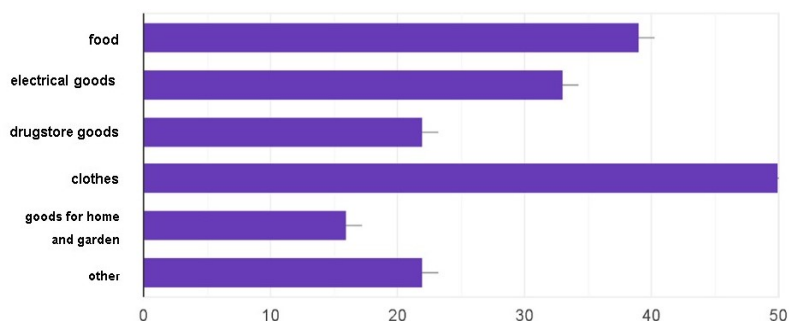
Figure 2. Overview of on-line shopping purchases during coronavirus period

When asked about what services offered by grocery and retail stores they value most, more than 40% responded that it is optimization of e-commerce and customer/channel strategy. More than 30% said that they value permanent behavior shifts to online and delivery across customer segments, shift for older consumers (50+) ordering on-line.

When asked about their opinion of consumers adopting short-term behaviors during the pandemic, many of which may become permanent, 62.5% believe that the biggest change will be a general switch to online shopping (Figure 3).

Question: Do you agree that consumer will have adopted short –term behavior during the coronavirus pandemic that in many cases will become permanent?
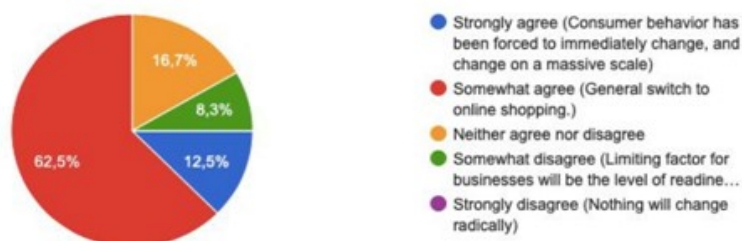
Figure 3. Understanding the short-term behavior and long-term consequences

We also asked our respondents if they believe that consumers will emerge from the pandemic in a new economic reality, changing commerce behaviors in profound ways. Almost 75 % of our respondents strongly or somewhat agree, and only 8, 3% somewhat disagree.

Question: Throughout history the spread of disease and economic crisis has been linked to lower emission. Do you agree, that coronavirus – COVID – 19 will have a lasting impact on a carbon emission?
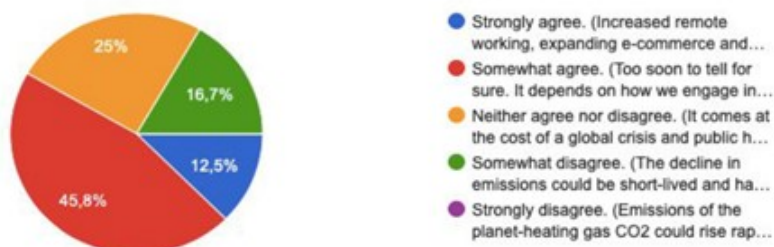
Figure 4. Overview of lasting impact of coronavirus-COVID-19 on carbon emission

192

Figure 4 above shows, that 45.8% somewhat agree but they think that it is too soon to tell for sure. Almost 16.7% of respondents somewhat disagree because their opinion is that the decline in emissions could be short-lived and have little impact on the concentrations of carbon dioxide.

4. Discussion and Conclusion

The results of primary research suggest with high probability that some habits and consumer behaviors developed over the quarantine coronavirus - COVID-19 period will be adopted by citizens when things eventually return to normal. It is highly likely that during the economic recovery people will continue to show the new behavior patterns, namely: saving money and avoiding spending money frivolously; shopping more cost consciously and buying more sustainable options; preferring to buy more product for health and well-being. It could be predicted over this period the less impulse buying or over buying being done by consumers and increasing shift towards proactive consumer behavior. These assumptions are in linked with Nejlsen study (2020) of 6 stages of behaviors during coronavirus, and are the subject of further discussion and research.

Conclusion. The presented study in a number of directions confirmed and specified some general tendencies that could be expected in purchasing behavior with the advent of the new millennium. The research findings confirmed, that:

- The popularity of payment via payment cards will continue to grow, becoming the dominant way of payment for purchase in nearest future;

- Today and in a nearly future the share of e-shop purchases will continue to increase, and the segmentation will undoubtedly increase purchases of groceries, ready meals and medicines over the internet. It is likely consumer behavior will change after the quarantine period generally shifting to online shopping with the increasing food purchases via e-shops. Manufacturers and distributors have to accept that consumers are much more well versed in goods comparisons, as more than a third of respondents and almost half of respondents are more likely to use price comparators to buy less expensive goods;

- During the coronavirus period an important behavioral trend in sustainable consumption was manifested. The consumers preferred to buy products and services that are healthy, trustworthy, environmentally sensitive, and supportive by the communities in which they live and work. This behavioral contextual reaction manifesting from a stress under pandemic may be long-term lasted behavior. This should induce forward-thinking businesses and marketers to leverage sensibility towards sustainability aspects, developing the circular economy and environmentally responsible green marketing.

- "Buy local" trend accelerates. Consumers shifting their spending's habits and some companies and product categories are actually growing. The native Czech brands Krasa, Rochlik, Košik are becoming increasingly popular among Czechs. Among the younger generation a positive trend in purchases of domestic food produced can be expected.

- It is worth mentioning that our represented group of younger generations, mostly Gen Z, does not mind schools during coronavirus period, moreover they continue to study hard, adapting quickly to distance on-line learning, what is the evidence of acceleration of society's digital transformation.

New patterns of consumer behaviours can serve as indicators for manufacturers, retailers and marketers. They can help companies adapt their supply chain to the shifting buying habits, re-evaluate their business models and better appeal to customers, encouraging them to purchase their goods or services.

5. References

Accenture (2020). Navigating the impact of COVID-19 in Consumer Goods, April 4, 2020

Amos, C., Holmes, G.R., & Keneson, W.C. (2014). A Meta-analysis of consumer impulsive buying. Journal of Retailing and Consumer Services, 21(2), 84-97.

Francis, T., & Hoefel, F. (2018). 'True Gen': Generation Z and its implications for companies. Article. McKinsey& Co, November,12, from:https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/true-gen-generation-z-and-its-implications-for-companies.

Espinoza, C., & Ukleja, M. (2016). Managing the Millennials: Discover the core competencies for managing today today's workforce. Second Edition, Wiley, NJ, USA, 2016, ISBN: 978-1-119-26168-1.

Mozer, Andrea K. Thinking green, buying green? Drivers of pro-environmental purchasing behavior. Journal of Consumer Marketing 32 (3), 167- 175.

Nejlsen Company (2020). Nejlsen Six Consumer Behavior Thresholds of COVID-19. Article, March 3, 2020, US, from: https://www.nielsen.com/us/en/insights/article/2020/key-consumer-behavior-thresholds.

Poncin, I., & Mimoun, M.B. (2104)."The impact of "e-atmospherics" on physical stores. Journal of Retailing and Consumer Services, 21(5), 851-856.

Ritzer, G., & Jurgenson, N. (2010). "Production, consumption, prosumption: the nature of capitalism in the age of the digital 'prosumer". Journal of Consumer Culture, 10 (1), 13-36.

Sinus-Milieus, (2018). Researching the right target market with Sinus-Milieus® Institute. November 2, 2018, Heidelberg, Germany, from: https://www.ionos.com/startupguide/productivity/sinus-milieu/

Slovenská Spořitelná CZ (2020). Household expenditures fell by almost a quarter. Instant Comment, April 6, 2020.

Statista (2020). "E-commerce share of total retail sales worldwide 2014-2023", August 30,2020

Švecová, L., Ostapenko, G. & Veber J. (2019) "Circular Economy in the Czech Republic". Proceedings of the 2nd International Conference on Education Science and Social Development (ESSD 2019). Advances in Social Science, Education and Humanities Research, Atlantis Press, July, 2019, 551-554.

Švecová, L., Ostapenko, G., Veber, J., & Valeeva,Y. (2019). The Implementation Challenges of Zero Carbon and Zero Waste Approaches. E3S Web of Conferences Journal,124 (October 25), doi:10.1051/e3sconf/201912404025.

Trudel, R., & Kotte, N. (2009). Does it pay to be good? MIT Sloan Management Review 50 (2), 61- 68.

SUCCESSFUL RECRUITMENT OF TEST PERSONS: A SYSTEMATIC RECRUITMENT STRATEGY

Georg Aumayr, Nadine Sturm, Gabriele Salomon, Gudrun Haider

Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH

georg.aumayr@johanniter.at

Keywords

Recruitment strategy, test persons, active and assisted living (AAL), active and healthy ageing (AHA), technology, innovation, evaluation

Abstract

Recruitment processes present end-user partners of active and assisted living (AAL) and active and healthy ageing (AHA) projects with diverse challenges. Involvement of people aged 65+as test persons is crucial for the development of digital solutions or technical devices. For this paper recruitment phases of former AAL and AHA projects including pilot sites for technology testing run by Johanniter Austria Research and Innovation (Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH, JOAFG) have been analysed. A systematic recruitment strategy was developed by JOAFG as guideline for recruitment processes and implemented to the current AAL project CARUcares. The strategy emphasizes the importance of activities tailored to the specific target group, precise defined networks including possible benefits of the technology in test and calculating sufficient time and personal resources.

1. Introduction

Numerous strategies and policies were developed over the last decade fostering digitalisation in various sectors. Even more, the European commission puts priority in digital development. Especially for devices used in every-day life, end-users are put at the centre of the developing process. Therefore the policy "Digital Inclusion for a better EU society" (European Commission, 2019) has been published whereby one focus is on assistive technology for elderly with the Active and Assisted Living Programme. Whether a technological solution is successful or not depends on acceptance, usability and impact for the defined user group. User involvement is mandatory at each stage of development to reach a high level of practicability. The term "end-user group" describes a certain group of people who are expected to benefit most from the technology developed. AAL stands for active and assisted living and subsumes technologies that improve autonomous living of older people at home. AHA subsumes projects dealing with active and healthy ageing. With a broad view of the living environment and conditions of older people, the concept of primary, secondary and tertiary end-user has been established. People 65+ are primary end-users, their closely related persons are secondary end-users and tertiary end-users include institutions and further organisations (Aumayr et al, 2017, p. 3f.). In recent years, numerous projects have been concerned with the development and evaluation of AAL technologies. User involvement varies, depending on the project objective and technology tested. However, to motivate persons to test a technology, different factors such as social constructs, culture, organizational frameworks, and personal interests must be taken into account. Participation in a technology-related study or research project often goes along with a confrontation

with unknown equipment or processes. This paper reflects upon experiences of previous recruitment phases of AAL and AHA projects and identifies key factors. Based on this findings a systematic recruitment strategy has been established.

2. Experiences gained in former projects

Over the past years, JOAFG run different pilot sites in various AAL and AHA projects. JOAFG was responsible for the recruitment and maintenance of regional pilot sites, mainly located in Austria, Vienna. Although all projects are related to the same topic they differ on various levels.

The project My-AHA (My Active and Healthy Ageing, Horizon 2020, 2016 - 2019) focused on frailty prevention. Main aim was the reduction of frailty risks in a multidimensional approach including cognitive, physical and social dimensions. Due to the project aims inclusion criteria to the 12 month randomized controlled trial (RCT) were very specific. The concept covers technology supported interventions on cognitive, physical and social levels. Therefore, 133 potential test persons were screened following Fried's frailty criteria (Fried et. al., 2010) to define their frailty status, which needed to be "pre-frail". Only 43 screened persons met the inclusion criteria and became test persons of the project (www.activeageing.unito.it). The project WAALTeR (Viennese AAL test region, national funding program in AT: benefit, 2016 - 2019) set focus on social integration and participation on urban life of older people. At the heart of the project stands an easy-to-use tablet application and a mobile emergency watch which were tested in a 12 months RCT with 137 test persons in Vienna. In a second phase of the project, participants could decide whether to test additional technologies in the field of security or e-health (www.waalter.wien). The project i-evAALution (AAL JP - EU, benefit - AT) evaluates smart home technologies, a tablet plus a mobile emergency watch as a combined and interactive bundle. A special constitution of this project is the evaluation of the impact of technology on informal secondary end-users (e.g. relatives). 45 people in Vienna are actually participating in a 12 months RCT (www.i-evaalution.eu). The project CARUcares (AAL JP - EU, benefit – AT, 2019 – 2022) is based on a voice controlled device developed in Switzerland by CARU AG. The aim of the project is to develop new functionalities especially for the sector of mobile care. On the one hand communication between older end users and their care persons or the care management should be facilitated. On the other hand, care persons will test care documentation functionalities. The project will contain two piloting waves, starting from 2021 on (www.carucares.com).

Testing digital solutions and innovative technologies in private contexts requires sensitivity and consideration of a range of ethical aspects. Implementing AAL solutions interferes with daily live routines of older people and goes along with confrontation with new devices which are still under development, as well as sensitive and intimate topics.

The role of end-user partners like JOAFG within AAL and AHA projects is to mediate between technical development and (primary) end-users. This emphasizes the importance of holistic approach in terms of digitalisation and innovation. From the very beginning, societal, psychological and sociological aspects have to be taken into account for developing new technologies as well as for testing phases. In terms of life-long learning self-efficacy, relevance and viability mainly determine motivation. Furthermore, factors like relevance and duration, interest and learning content, social acceptance and educational added value influences the perception as a learning individual. In addition, the own biographical background plays a significant role regarding motivation and learning behaviour (Siebert, 2008, p. 180ff.). Therefore, educational, cultural and social determinants have to be taken into account. Motivation is expected to be higher when people get the impression to be fully involved and their opinions and concerns are taken seriously (Lorig, 2001, p. 144ff.).

2.1. Recruitment processes in health and technology related projects

Recruitment strategies are defined for each project in line with the project plan. Although communication measures were taken and network activities conducted, it is one of the main findings that it is hard work to attract (especially older) persons to finally participate on a study. The figure below shows schematically the relationship between initially interested persons and real study participants for the above described projects.
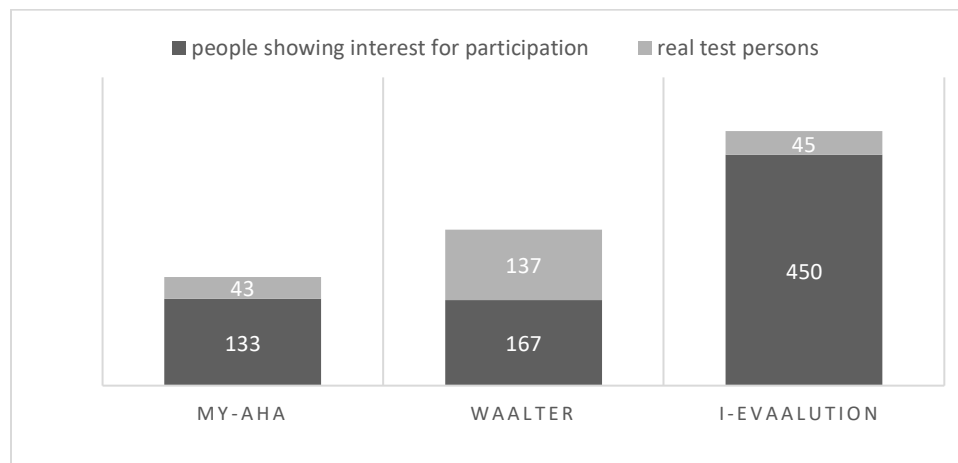
Figure 1 Recruitment numbers, Johanniter 2020

2.1.1. Perceptions about recruitment activities

Although a huge effort is taken to attract the target group to take part on a study the outcome number of test persons seems to be low in general. Many factors influence this situation, ranging from scepticism and fear of technologies, high efforts for test persons, too complicated study designs or the possibility to be randomized and allocated to the control group. Following key finding could be identified:

- Activities with wide reaching potential, like TV spots or newspaper advertisements, contribute to awareness (e.g. on digitalisation or technology) but are too unspecific to make a decision whether to participate or not.

- At face to face information events more detail information can be provided and individual questions answered. Another positive effect is the possibility to demonstrate the respective technology on site. Small groups enable better response to individual questions and concerns.

- A combination of general, wide reaching information and information event could be a sufficient way.

- Calculate resources for network activities. It is likely that access to potential primary end-users via potential secondary end-users is more effective.

Potential participants must get transparent information to make a self-determined and informed decision. This information must not only cover technology related issues but specifics about the study design. This is even more important when stepping into a randomized controlled trial with the possibility to be allocated to the control group.

2.2. Systematic recruitment strategy for AAL projects

Based on the findings and experience in previously described AAL and health related projects a systematic recruitment strategy has been developed. The aim is to provide a systematic step by step

process as well as foster visualisation. This enables the recruiter to plan tasks, get an overview of necessary steps and detect gaps in a very early stage. It is recommended to go through the below described steps before starting the recruitment process. The presented recruitment strategy is actually used for the project CARUcares (Sturm et al, 2020, p. 13ff.).

The systematic strategy is composed of three pillars:

- Reflecting on key questions
- Planning further steps using visualisation tools
- Using a checklist to ensure every important detail is covered

2.3.   Reflecting on key questions

The first step is to reflect some key questions which contain determining factors of the recruitment process. Answering these questions states the overall situation and further steps can be defined on that basis.

- Which end-user groups shall be involved? What are the inclusion/exclusion criteria for test persons? How many test persons are required per end-user group according to application?
- Which structure can be defined, in case of organizational involvement?
- Who will test which functionality/technology/devices? At which development status will test persons get the test equipment?
- What is the main benefit for whom?
- How can potential participants be reached? Which communication channels shall be chosen? Which information material is suitable?
- Which resources are required (staff, working hours, other resources)?
- Are there any additional or specific circumstances to be considered?

2.4.   Planning further steps using visualisation tools

For detailed planning of further steps, the following visualization tools are suggested (Sturm et al, 2020, p. 13ff.). First of all, the target groups need to be identified and main benefits per end-user group of the technology in test have to be stated. This assumptions are based on the researcher's ability to consider the specific end-users living situation, the context the technology will be tested in and how it will ideally contribute to enhancement of quality of live in different dimensions. It is recommended to imagine the values (and thus requirements) for a customer for the product if it would already be available on the market. Therefore, also consider the system of Business Model Canvas and its included value proposition (for further information visit: www.strategyzer.com).

Primary end-user (e.g. older person) Insert benefits:	**Secondary end-user** (e.g. relatives): Insert benefits:	**Tertiary end-user** (e.g. care organization): Insert benefits:
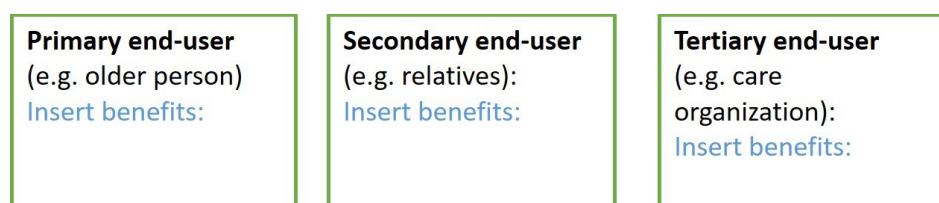
Figure 2 Main benefits per end-user group, Johanniter 2020

For the planning of next recruitment steps it is necessary to be aware about envisaged numbers of test persons in a project application or by another contracting authority. Furthermore, different phases

of end-user involvement into the project according to the project plan need to be considered. It is helpful to include the research setting, as it makes a huge difference if test persons shall be recruited for focus groups or large trial phases. Additionally the duration of the phase can be included. The figure below shows an example to illustrate the respective numbers.

Primary end-user (e.g. older person)	Secondary end-user (e.g. relatives of primary end-user)	Tertiary end-user (e.g. care organization)
Phase 1 Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 1: insert	**Phase 1** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 1: insert	**Phase 1** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 1: insert
Phase 2 Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 2: insert	**Phase 2** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 2: insert	**Phase 2** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 2: insert
Phase 3 Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 3: insert	**Phase 3** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 3: insert	**Phase 3** Required number of test persons: insert Research setting: (e.g. field trial) Duration of phase 3: insert

Figure 3 Required number of test persons per end-user group, Johanniter 2020

To define appropriate communication channels it is necessary to consider, how far the scope of a communication medium must be. The following graph differentiates between communication channels with wide and narrow reach. It needs to be decided according to characteristics of the target group as well as numbers of test persons required which method of communication fits best.
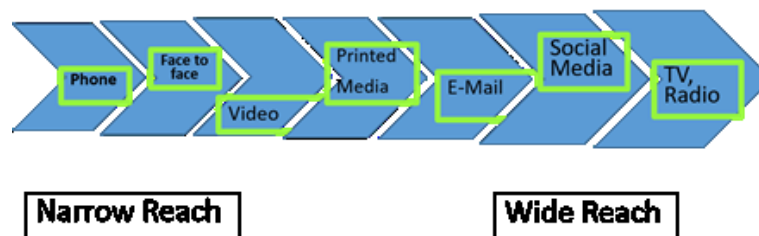
Figure 4 Communication channels, Johanniter 2020

The figure below illustrates how communication paths could look like. It is important to define, how a certain end user group can be reached best and who the ideal communicator is. Depending on the preconditions of the research field information brochures, press releases, social media postings or information events etc. can be created.
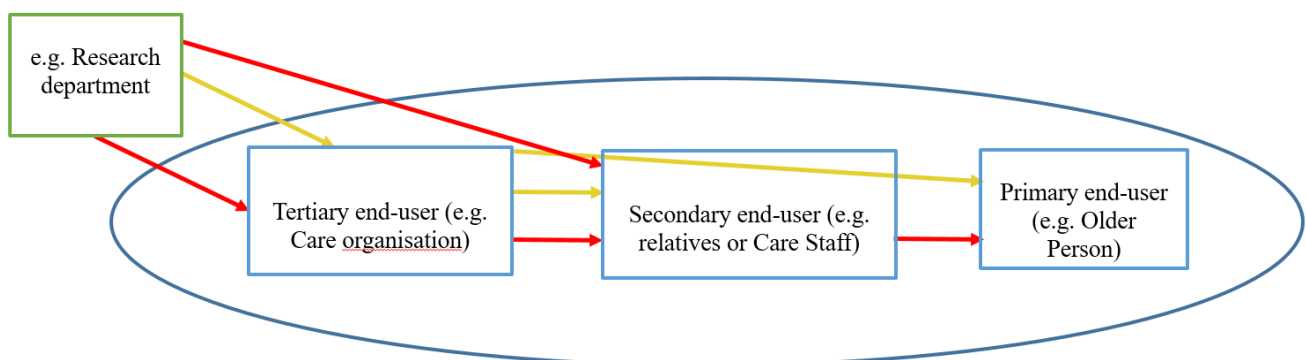
Figure 5 Communication paths, Johanniter 2020

In the last step the timeline for recruitment measures is defined. As pretesting often needs to be finished before starting testing phases it is recommended to consider this phase parallel to recruitment activities. The timeline aims at providing a rough overview on testing phases but can be enlarged to plan specific recruitment activities for the target groups in a specific timeframe.

PRE TESTING by insert organization name Team:
- Insert timepoint/timeframe
- Insert other option

Primary end-user:
- Phase 1: finished
- Phase 2+3: Insert timepoint/timeframe

Secondary end-user:
- Phase 1: finished
- Phase 2+3: Insert timepoint/timeframe

Tertiary end-user:
- Phase 1: finished
- Phase 2+3: Insert timepoint/timeframe
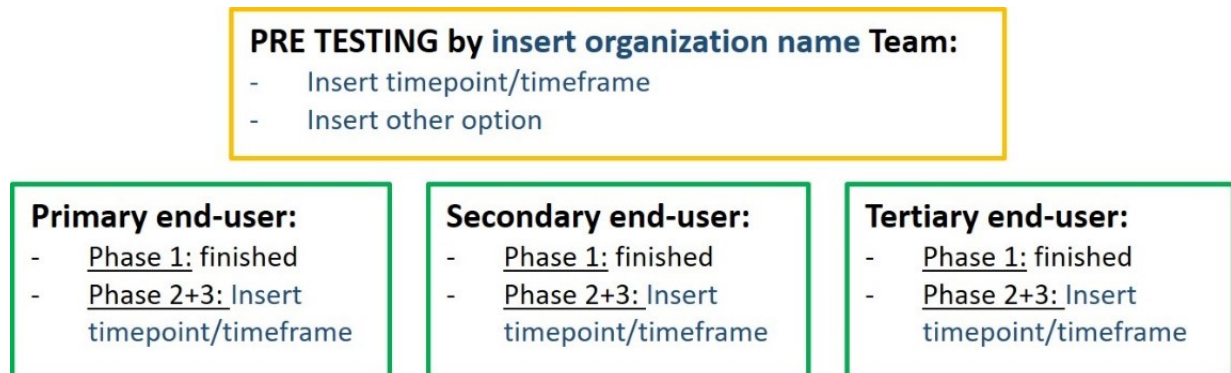
Figure 6 Required information material per end-user group, Johanniter 2020

Following these guidelines will enable planning and conducting recruitment according to predefined timelines and project aims. It will further contribute to detect gaps at a very early timepoint and take appropriate countermeasures. It is recommended to create a checklist and enlarge key questions as well as to adapt this recruitment planning strategy in line with the respective project plan.

3. Conclusion

Johanniter Austria Research and Innovation (Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH) took part in several AAL (active and assisted living) and AHA (active and healthy ageing) projects over the past years. As an end-user partner, Johanniter had the responsibility for several pilot sites for technology testing. Nevertheless, recruitment of end-users to participate in the projects as test persons is always a crucial and challenging task. Thus, the recruitment process of former projects have been reflected. Key issues that have been identified, which emphasize the importance of activities tailored to the specific target group, precise definition of the target group and calculating sufficient time and personal resources. The relationship between project team and end-user has to be understood as a bilateral and dynamic process at every stage. Based on the results a systematic recruitment strategy was developed and integrated to the current AAL project CARUcares. This strategy contains key question and a systematic path through a recruitment planning phase. It aims at visualizing major steps to be able to identify gaps at a very early stage. A timeline is provided which supports keeping deadlines and meeting recruitment numbers. This systematic recruitment strategy is meant to support other end user partners and is used by Johanniter to plan recruitment activities in AAL and AHA projects.

To understand the needs and wishes of end-users is essential for the path of technology into living practice. However, according to international policies like "Digital Inclusion for a better EU society" of the European Commission, more multidisciplinary research about involvement of participants is necessary.

4. References

Aumayr, G., Bleier, D., Sturm, N. (2017). Requirements and Pitfalls in AAL projects. Guide to Self-Criticism for Developers from Experience. Multidisciplinary Digital Publishing Institute (MDPI): Informatics-Basel (Vo. 4, No. 42).

CARUcares, (2020). Retrieved from www.carucares.com, 16th May 2020

European Commission (2020). A Europe fit for the digital age. Empowering people with a new generation of technologies. Retrieved from https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_en#documents, 16th May 2020

Fried, L., Tangen, C.M., Walston, J., Newman, A., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W., Burke, G., McBurnie, M., (2001). Frailty in Older Adults: Evidence for a Phenotype. Journal of Gerontology: Medical Sciences, Vol. 56A, No.3, M146-M156

i-evAALution, (2020). Retrieved from www.i-evaalution.eu, 16th May 2020

Lorig, K. (2001). Patient Education. A Practical Approach (Vol. 3). Thousand Oaks, California, USA: Sage Publication, Inc.

My-AHA, (2020). Retrieved from www.activeageing.unito.it, 16th May 2020

Siebert, H. (2008). Konstruktivistisch lehren und lernen. (R. Feuchthofen, M. Jagenlauf, & A. Kaiser, Eds.) Augsburg: ZIEL-Zentrum für interdisziplinäres erfahrungsorientiertes Lernen GmbH.

Sturm, N. (2020). Strategies for user recruitment, Del. 5.1. Caru Care Project consortium , Vienna.

WAALTeR (2020). Retrieved from www.waalter.wien, 16th May 2020

SOCIAL MEDIA, PRIVACY AND ETHICS

PRIVACY ON SOCIAL MEDIA – SHARING INFORMATION ON DIFFERENT SOCIAL NETWORKS

Antonín Pavlíček, Jana Syrovátková

Department of System Analysis

Faculty of Informatics and Statistics

University of Economics Prague

Keywords

Social media, privacy, information sharing, LinkedIn, Facebook, Twitter, Instagram

Abstract

The paper compares information people share on social networks and analyze the differences. Data were acquired from public profiles (N=200) on LinkedIn, focusing on young ICT specialists, comparing with the profile content on other social networks - Facebook, Instagram and Twitter. We managed to get enough data to perform some basic statistical analysis to compare with our hypothesis. Conclusions: we have managed to track more than a half of the profiles on Facebook and over a quarter on Instagram and Twitter. People were quite open on Facebook, almost 80% of them sharing at least some personal information but only 23% on Instagram. The paper offers an insight to the approach people take when sharing information on social networks and keeping their privacy safe.

1. Introduction

1.1. What people share on social networks

There is increasing amount of data and information entities that are of significant importance for many institutions, businesses (Klimek et al, 2017) and even citizens as the part of the digital transformation within many fields in our society. (Cerny, Potancok, Molnar, 2019) Social media became a natural communication tool used by generation Y. Young people prefer to use social media as the primal system which provide them the functions for managing their personal agendas. (Vondra, 2017) At the same time, people consider online privacy more and more important. However, privacy is a complex issue, which can be treated in many perspectives. (Sigumund, 2017) Users' privacy and security is important in terms of sharing information on social networks. According to study by Ipsos and Reuters, only about 75 % users understand privacy settings. The rest does not care, or does not know how to change privacy settings. Currently, fewer users share their posts publicly - they tend to have a private (friends-only) profile. Study indicates that the number of publicly shared contributions keeps decreasing. On the other hand, only 23% of users believe that they have absolute control over the privacy of their contributions. (Nemeckova, 2018)

The issue of willingness to openly share information can be measured by the question of whether users avoid sharing controversial content on Facebook. The results are quite clear – respondents answered: completely true 21 %, true 31 %, neither true nor untrue 16 %, untrue 18 % and completely untrue 14 %, which shows that this issue is evenly spread throughout the population. (Hutchinson, 2016)

Research from the University Brunel in London points out that, according to shared contributions on social networks, it is also possible to identify a person's personality. Their research was based on a five-factor personality model called "Big Five". This model, which is fairly popular, determines five dimensions of human personality: According to the ''Big Five'' model of personality, individuals vary in terms of extraversion, neuroticism, openness to experience, agreeableness, and conscientiousness. (Marshall et al., 2015). Big Five model is quite popular and can be used for various purposes, e.g. Impact of Gender and Personality Traits (BFI-10) on Innovativeness (Pavlicek, Sudzina, 2017a), or Impact of gender and personality traits (BFI-10) on trust (Pavlíček, Sudzina, 2017b) or Personalities in the Preferred Pricing of Public Transport (Pavlíček, Sudzina 2020).

Fractl's research studies the main reasons for sharing social media posts. These reasons basically create the basis for what people share. According to the research, 48 % of users share their contributions just for fun and to amuse their friends. Another 17 % of users share posts to express their attitudes and concerns. Another reason is "Move" - posts, which will make friends feel something. It means feelings like happiness, sadness, anger and so on. This reason is represented by 13 % of users. An interesting reason for sharing posts is education, which is represented by 11 % of users. The remaining 11 % of users represent other (not specified) reasons. (Hutchinson, 2016)

The 2014 survey from the book "Social Media in Employee Selection and Recruitment" showed, that 51 % of employers gets rejected from job applicant based on some problematic information found on social sites. The author of the book also states that employers are well considering content of their social networks and tactics to reach candidates. (Richard at al., p. 2-3, 2016)

Böhmová and Novák (2015) have proved, that LinkedIn is a very important recruiting tool for companies by comparing how the ten biggest companies in three different European countries (Norway, France and Germany) used LinkedIn for hiring in comparison with the most popular job boards. Their overall finding however is that there still was potential to grow, since LinkedIn has not been used as much by companies as theory suggests it should have been.

We should also mention the research from Slovakia. The participants in their surveys said that social networks as a tool for job search are perceived as less trustworthy – this may be the cause of a low number of job offer responses. This option was indicated by 58.77 % of the Slovak and 54.14 % of the foreign students and university graduates. They recommend recruiters in Slovakia to focus on Facebook, Pokec, Skype and YouTube. But they also mention other research carried out by LMC company, which considers most effective Facebook and LinkedIn, because there are used for work related issues by 24 % and for personal use by 27 % of the 234 respondents. (Bartakova et al., 2017)

Other research showed that in 2013, 56 % of employers used social media for recruiting. Today that number has risen to 77 %, according to a recent survey. According to their research LinkedIn is by far the most popular website for HR professionals engaged in recruiting, with 94 % active users. Facebook trails with 54 %, Twitter follows with 39 %. (Lejeune, 2013)

And another research focused on software industry (data collected in Sri Lanka) indicates that 80 % of the HR professionals use LinkedIn, 49 % use Facebook, 12 % use Twitter, 2 % use YouTube, and 22 % do not use any of these social networks. (Amadoru, Gamage, 2016)

These findings are confirmed by research from the Czech Republic. Results found that the number of recruiters using Social Networking Sites to publish job posts in search of new employees doubled compared to the year 2012. In 2012 it was 30,45 %, now (in 2016) it is 60 %. Also claims that the most frequently used platform for recruitment was LinkedIn. (Bohmova at al., 2015)

In our paper, we have decided to evaluate how people protect their data on social networks and what type of information do they share. Then compare the differences in shared data on various social networks – in particular LinkedIn, Facebook, Twitter and Instagram. LinkedIn is social network where people build their work portfolio and get job offers, so it would be reasonable to expect for

them to share information about their skills and previous job experience. On the other hand, on Facebook, Twitter and Instagram we are expecting more informal personal information.

2. Methodology

During our research, we tried to get as much information about different people as possible. The core of the goal is to find them on more than one social network and get the whole picture. Later, we have approached them with the same offer via different social networks and evaluate their responses based on the platform used. From that, we should find out what way is the best and most effective for recruiters to communicate with strangers, possible prospects. One of the components of the research is to find out if the people contacted even connect with us, letting us see more personal information about them even if we are strangers.

The research was conducted on the basis of the following research methods:

- Analysis of
    - profiles on social networks (LinkedIn, Facebook, Twitter, Instagram),
    - people's security settings on social networks,
    - profile findability and credibility,
- Contacting people on social networks,
- Observation of reactions,
- Comparison between the networks.

2.1. Hypotheses

We came up with following hypotheses, which we considered to be possible and interesting to investigate:

1. People, who have more than one photo on their profile, are more open and tend to share more information.
2. People with profile full of information on Linked will also share more stuff on other sites.
3. Profiles with more information are also more active.
4. People in the analysed set of data generally have a lower followers-to-following ratio than Twitter and Instagram users in general.
5. People will react on job offers more on LinkedIn than Facebook.

2.2. Data gathering

Our sample consisted of people who have profile on LinkedIn and are open to new job opportunities. They are IT workers - especially programmers from Brno city.

2.2.1. Linked In

We have been looking for following attributes in all the profiles: "Name" (first and family name), "gender" (M or F). Next observed fact was "LinkedIn:foto" which consist number of photos, which we can find on LinkedIn profile in a specific person. "LinkedIn:city" means town, in which person live. Information "LinkedIn:schools" consist statement if person shares information about schools

where he or she studied, we judge especially high schools. "LinkedIn:act_position" - information whether there is on LinkedIn profile published name of actual position and name of company where they work. With relation to this, there is also "LinkedIn:prev_position" and there we looked on previous positions. "LinkedIn:languages" gathers information about languages people persons speak. "LinkedIn:projects" means that persons share information about some previous work-related projects on LinkedIn. "LinkedIn:skills" is about knowledge which people have, it is for example some programming languages (Java, JavaScript, C#, C, PHP etc.) and we looked if this information are shared. We looked also on visible groups, which persons likes - we tagged it as "LinkedIn:groups". Alike information is also "LinkedIn:following", it is interest and fun groups, which people likes and watch continuously, they make there special act to have actual posts from this site in their profile wall. "LinkedIn:recommendation" is about fact, that on LinkedIn you can require for a recommendation from your boss or coworker. We examined if there is information about it. "LinkedIn:pers_web" it is about shared personal web, if person share his or her website. This information we collect also about email (LinkedIn:email). "LinkedIn:additional_info" it is additional information - if on LinkedIn profile have people other self written personal information (for example some status, other work experiences, other way to contact them and others). Some people share and give attachment on their profile - it is often Curriculum Vitae - "LinkedIn:CV_attach". We examine also their birthday (LinkedIn:birthday). Sometimes people share contact to their other social sites, on LinkedIn you can share your Twitter profile - "LinkedIn:Twitter" - there you can find nickname of people who share it. Finally with LinkedIn data, there are numbers of connections, which people have witch other people - "LinkedIn:connections, it works like friends on Facebook, people can be connected there with friend, coworkers, recruiters or business partners.

For every subject, we have maintained also "LinkedIn:URL_profile" – full address of the observed LinkedIn profile.

2.2.2. Facebook

This social site have been searched for information regarding users from already analyzed LinkedIn profiles. Depending on privacy settings, we have noted whether the profile was secure or public. When the profile is secure, there is not most of the information like friends, hobbies, cities, family relationships, or wall posts = therefore we have recorded attribute "FB:secure". For public profiles have been recorded following attributes: "FB:photo_number", where is number of photos sorted into groups, and "FB:photos" which is the result of our inclusion in small x medium x large number of photos. The following attribute was "FB:information", which means that the person in the information section has some shared data. "FB:birthday" and "FB:born_town" if birthday and bithrplace were public. "FB:live_in" information about actually living of person. We examined whether a person shares data about their current work in "FB:current_work", about schools they attended "FB:school". "FB:relationship" means in which relationship person is. We have also looked into the information about their friends - "FB:friends". Next information is "FB:family", where we recorded "0" or "1" - "0" when person do not share information about their family members, and "1" where we can find their family members - they are connected and emphasized as family members. And "FB:like_pages" is about rate of security with pages which person likes. We marked "1" if we can see any pages which are favourite for person. Finally, there is "FB:main_content", where we search for content of person profile, we divided into categories: events, sport, family, friends, travel, education, selfie, TV, music, books, work, food, entertainment, politics.

2.2.3. Twitter

Social network Twitter was searched in the same way as Facebook, using LinkedIn user list. We have determined whether the person is active on twitter. We have recorded attributes: "Twitter:following"

as a number, how much people or sites the person is watching on Twitter. And "Twitter:followers" as the number of people watching observed person. Next recorded information was about personal photo (with face). Like in Facebook, we also watched and wrote the main content of their posts on Twitter. We divided it into the same categories as Facebook in "Twitter:content".

2.2.4. Instagram

Instagram profiles were searched based on LinkedIn, Facebook and Twitter information. For each Instagram account was recorded a nickname - "IG:nickname". As well as in other social networks, we have evaluated public/private setting under "IG:secure". "IG:profile photo" was recorded, when a person has a profile photo with a face. Also, if he has written some information about himself (nation, city, interest, age, etc.) - "IG:information" was filled in. We have also recorded number of posts on profile wall - "IG:posts". Like in Twitter we have monitored number of followers "IG:followers" and number of following - "IG:following". "IG:content" noted areas of interest based on posts on profile wall, in the same structure as in Facebook.

3. Results

In this chapter our most interesting analysis are described. We started with 205 profiles on LinkedIn website and matched them to 121 Facebook accounts, 81 Twitters and 57 Instagram accounts. This shows that it is relatively easy to connect particular person's profile to another on different social network. On Facebook there are formulated question about personal information and the profiles are much more complex than on Instagram. This makes it easier to find specific people and that is the reason behind the number difference.

3.1. How easily you can get personal information

3.1.1. Amount and share of private accounts

The hypothesis for this part was that people tend to make their Instagram profile private more than Facebook, because it's easier and the motivation might be slightly higher due to the content being just photos. We also expected that the share of private profiles wouldn't be higher than 30% and that turned out to be confirmed.

The results were following:

Table 1: Amount and share of private accounts (source: authors)

Network	Private accounts	Total	Share in %
Facebook	28	121	23 %
Instagram	17	58	29 %

3.1.2. How often people share basic information

On LinkedIn only 7 people out of 206 did not share information about their current job position or school they went to. This is obviously due to the fact that this is the most important type of information people should share on LinkedIn because of its purpose. Often they also add languages skills and any other job related information to make better first impression on possible future employer. But what about other - not so serious social networks?

Starting with Facebook, almost everyone filled (or decided to share) at least one of the personal information asked by Facebook sign-in forms, which includes for example date and place of birth. This information can be useful when searching for people from your past or even present, because it adds additional keywords to just name. That is the reason why we weren't surprised by such big number people sharing this information. On the other hand, on Instagram there is no such feature and the information to share are only in plain text are so the number is much lower.

Table 2: How often people share basic information (source: authors)

Network	Number of accounts	Information shared	Percent
Facebook	121	95	79%
Instagram	57	13	22%

3.2. How often do we share

There seems to be relation between sharing basic information and number of photos on Facebook. If you are being open about your personal information, chances are you are also sharing more photos. After checking every profile gallery, we classified the number of photos on each profile into groups - low (0-10 photos), medium (11-50 photos) and high (51+) and run correlation. The results were not far from expectations. Almost 67 % of profiles without any information had low number of photos and 25 % medium. On the other hand, between the profiles with at least some information, the numbers were much more equal - 40 % had low and medium, 20 % high number of photos (see the table):

Table 3: How often do we share (source: authors)

Number of photos	Profiles with information: 96	Without information: 25
low	38 (40 %)	16 (67 %)
medium	38 (40 %)	6 (25 %)
high	20 (20 %)	2 (8 %)

3.3. Formality of the content topic

Another objective of our research was to assess the behavior of various social media users in terms of the formality of their posts. We have used a list of 14 different categories of content people can share on their social media profiles and ranked them on a scale of 1 to 5 to assess the level of formality. For example, content about education or politics were considered to be the most formal and thus ranked 5, while selfies or sport related content were ranked 1. Based on this list, we have created an overview of the number of occurrences of different content categories for each of the three analyzed services - Facebook, Twitter and Instagram. Then we have summed up each respective category for each social media and weighted them to create a resulting index showing the level of formality for each of these services. The analysis found that on Twitter, people mostly post very formal content while on Instagram, the level of formality is significantly lower - this is evident just by looking at the highest rated category, which shows that almost 25% of Twitter posts are education or politics related while only 2,7% of Facebook posts and 0 Instagram posts fall into the same category.

Table 4: Formality of the content topic (source: authors)

Content type	Level of Formality	Occurrences			Weighted		
		FB	Twitter	Instagram	FB	Twitter	Instagram
education	5	2	14	0	2,7%	24,8%	0,0%
politics	5	2	11	0			
family	4	14	1	4	41,9%	37,6%	34,9%
travel	4	38	2	22			
books	4	1	3	0			
work	4	9	32	4			
friends	3	20	0	13	13,5%	0,0%	15,1%
events	2	16	2	4	22,3%	11,9%	20,9%
tv shows/films	2	5	5	0			
music	2	8	4	2			
food	2	4	1	12			
sport	1	16	6	11	19,6%	25,7%	29,1%
selfies	1	10	1	12			
entertainment	1	3	19	2			
Resulting Index					2,86	3,24	2,56

3.4. Followers ratio

"People in the analyzed set of data generally have a lower followers-to-following ratio than Twitter and Instagram users in general."

On two of the analyzed services - Twitter and Instagram, it is possible to view the statistics for both how many people a user is following and how many people the same user is followed by. Using these two statistics, it is possible to calculate the followers-to-following ratio, sometimes referred to as the follow ratio. The objective of this hypothesis is to determine, whether or not for the analyzed set of people this ratio is lower than in the general population. For this particular analysis we have used the results of two previous studies (Byrne, 2014; Beevolve, 2012) which show follower statistics on Instagram and Twitter. These studies show that on average the follow ratio on Twitter is 2.04 and 1.03 on Instagram. For our set of data, the Twitter follow ratio is 0.33 and 1.61 on Instagram. Even though the follow ratio in our set of data is indeed lower than in the general population of Twitter users, it's higher for Instagram users and therefore the hypothesis 5 was rejected.

3.5. Reply on job offers on LinkedIn and Facebook

As mentioned in the previous section, we have addressed people on their profiles on LinkedIn and Facebook. On LinkedIn, we send them request for connection and short message, where we wrote some job offer. When person accepts request for connection, it is sign that he or she has read the message. Next thing we looked into replies - if people send reply on our job offer. They can write refusal "thank you, I am not interest", or indicate "it is interesting, I want more information about this offer" and others ways. On Facebook, we tried to send them similar short message, and also

examine whether a connection has been established and person read a message, and if they send to us some reply on this job offer.

The only limitation is, that we have less Facebook profiles. The reason can be that some person do not have Facebook profile, somebody do not have profile photo and we can not find the relevant profile of the person, and sometimes on Facebook are more possible profiles which meet the criteria and we do not have 100% certainty. So, we have total 205 people, which we approached on LinkedIn, and reply was send to 116 people on Facebook - it is 57 %.

A total of 16 people on Facebook and 30 people on LinkedIn read our job offer message. It is clear that LinkedIn was more successful. Overall, 60% of respondents did not read our offer. On LinkedIn job offer was written by 26% of people. Facebook results were lower - only 14%.

And what about replies on job offers, which we send. We can see that 84% of the total number of approached people do not respond to the job offer. Only 9% addressed on LinkedIn replied to the job offer, and 7% on Facebook. Although the results are not satisfactory, we can conclude from both graphs that for recruiters has a much better chance to reach people on LinkedIn. It has been confirmed that this professional network is more prepared for job offers. The hypothesis "People will react on job offers more on LinkedIn than Facebook" is confirmed.

4. Conclusion

The conducted research is an interesting window into the thinking of young people and their activity on different social networks. We managed to track a sufficient amount of people across different social networks. One of the deal breakers was whether the profiles had profile picture or not, since this aspect was the most deciding when it comes to more profiles with the same name. Generally speaking, people tend to tidy their profiles on LinkedIn the most, because it might help them in their pursuit of greater career. On Facebook, they share some basic personal information, which is helpful when connecting with long lost friends or searching for someone they met recently. Topics of the content shared also differ, with Twitter getting the best grade from this perspective.

5. Limitations of the study

The research was conducted on a sample of just 205 users of the professional social network LinkedIn - specifically a sample of IT programmers from Brno. If the sample was more extend, then the results could be different. Secondly, sample contains only 8 women (due to focus on ICT industry), which could have skewed the results.

6. Acknowledgment

7. References

BARTAKOVA, G. P., BRTKOVA, J., GUBINIOVA, K., HITKA, M. (2017). Actual trends in the recruitment process at small and medium-sized enterprises with the use of social networking.[online] Economic Annals-XXI. Available on: http://soskin.info/userfiles/file/Economic-Annals-pdf/DOI/ea-V164-18.pdf

BEEVOLVE, (2012). An Exhaustive Study of Twitter Users Across the World [online] 2012-10-10 [cit. 2020-04-12]. Available on: http://www.beevolve.com/twitter-statistics/#b1

BÖHMOVÁ, L, NOVÁK, R. (2015). How Employers Use Linkedin for Hiring Employees in Comparison with Job Boards. In: IDIMT-2015 Information Technology and Society Interaction and Interdependence. Poděbrady, 09.09.2015 – 11.09.2015. Linz : Trauner Verlag Universität, 2015, s. 189–194. ISBN 978-3-99033-395-2.

BYRNE, M. (2014). Instagram statistics [online]. 2014-03-12 [cit. 2020-03-11]. Available on: https://opticalcortex.com/instagram-statistics/

ČERNÝ, J., POTANČOK, M., MOLNÁR, Z. (2019) Using open data and Google search data for competitive intelligence analysis. Journal of Intelligence Studies in Business JISIB [online]. 9/2, pp. 72–81. ISSN 2001-015X. available at: https://ojs.hh.se/index.php/JISIB/article/download/470/216.

DOUCEK, P., MARYŠKA, M., NOVOTNÝ, O. (2014) Requirements on the competence of ICT managers and their coverage by the educational system – experience in the Czech Republic. Journal of Business Economics and Management [online]. Vol. 15/5, pp. 1054–1077. ISSN 1611-1699. Available from: http://www.tandfonline.com/doi/abs/10.3846/16111699.2012.658436?journalCode=tbem20#.VH3j1MmdA4y.

HUTCHINSON, A. (2019). New Study Looks at What Motivates Facebook Users to Share Content [online]. SocialMediaToday, 2016-06-22. [cit. 2020-04-26]. Available on: https://www.socialmediatoday.com/social-business/new-study-looks-what-motivates-facebook-users-share-content-infographic

KLÍMEK, J., KUČERA, J., NEČASKÝ, M., CHLAPEK, D. (2018) Publication and usage of official Czech pension statistics Linked Open Data. Journal of Web Semantics. Vol. 48, pp. 1–21. ISSN 1570-8268 DOI: 10.1016/j.websem.2017.09.002. Available at: https://www.sciencedirect.com/science/article/pii/S1570826817300343

LEJEUNE, T. (2013) Use of Social Media to Recruit Increasing. Employee Benefit News, 2013. Available at: https://www.benefitnews.com/news/use-of-social-media-to-recruit-increasing-2732724

MALMI Amadoru, Chandana Gamage. Evaluating Effective Use of Social Networks for Recruitment [online]. Conference on Computers and People Research, 2016. [cit. 2019-05-09]. Retrieved from: https://dl-acm-org.zdroje.vse.cz/citation.cfm?id=2890604

MARSHALL, T. C., Lefringhausen, K, Ferenczi, N. The Big Five, self-esteem, and narcissism as predictors of the topics people write about in Facebook status updates [online]. Journal Personality and Individual Differences, Elsevier, Volume 85, pages 35-40, 2015. [cit. 2019-05-18]. Available on: https://www.sciencedirect.com/science/article/pii/S0191886915003025

NĚMEČKOVÁ, K. Studie: 74 % uživatelů stále používá Facebook každý den [online]. FOCUS agency s.r.o. 2018-5-16. [cit. 2019-05-19]. Available on: https://www.focus-age.cz/m-journal/aktuality/studie--74---uzivatelu-stale-pouziva-facebook-kazdy-den__s288x13699.html

PAVLICEK, A., & SUDZINA, F. (2017a). Impact of Gender and Personality Traits (BFI-10) on Innovativeness. In A. Kocourek (Ed.), Proceedings of the 13th International Conference: Liberec Economic Forum 2017 (pp. 234–239). Technical Univ Liberec, Faculty Economics., ISBN 978-80-7494-349-2, WOS:000426486500026

PAVLICEK, A., & SUDZINA, F. (2017b). Impact of Gender and Personality Traits (bfi-10) on Trust: Managerial Approach. In T. Loster & T. Pavelka (Eds.), 11th International Days of Statistics And Economics (Vol. 11, pp. 1163–1170). Melandrium. ISBN: 978-80-87990-12-4, WOS:000455325300116

PAVLICEK, A., & SUDZINA, F. (2020). Intergroup Comparison of Personalities in the Preferred Pricing of Public Transport in Rush Hours: Data Revisited. Sustainability, 12(12), 5162. https://doi.org/10.3390/su12125162

RICHARD, N. and collective. Social Media in Employee Selection and Recruitment: Theory, Practice, and Current Challenges. Springer, 2016. 376 pages. ISBN: 3319299891

SIGMUND, T. (2017) Model of Online Privacy. In: IDIMT-2017 Digitalization in Management, Society and Economy [online]. Poděbrady, 06.09.2017 – 08.09.2017. Linz : Trauner Verlag Universität, pp. 173–180. ISBN 978-3-99062-119-6. Available from: http://idimt.org/wp-content/uploads/proceedings/IDIMT_proceedings_2017.pdf.

VONDRA, Z. (2017). Initiative To Support Use of Social Media Functions to Improve Education. In: IDIMT-2017 Digitalization in Management, Society and Economy [online]. Poděbrady, 06.09.2017 – 08.09.2017. Linz : Trauner Verlag Universität, pp. 155–162. ISBN 978-3-99062-119-6. Available from: http://idimt.org/wp-content/uploads/proceedings/IDIMT_proceedings_2017.pdf.

HOW SCIENTIFIC YOUTUBERS TACKLE EVIDENCE-BASED APPROACH: A FULL STORY

David Anthony Procházka

david.prochazka.km@vse.cz

Keywords

Influencers, YouTubers, content makers, evidence-based practice, information sources

Abstract

Youtube influencers are in high regard in our society, and their numbers and numbers of their followers grow exponentially. However, are the information they share evidence-based? How do YouTubers tackle the validation of their sources, and is it important to them at all? Well-known YouTubers that claim to be working with scientifically proven information were approached in one-on-one semi-structured interviews in person or via Skype to find out how they look for and verify the information they use. The results are baffling as they show that most of the approached YouTubers know how they should look for data and critically appraise it, however not one has effectively implemented any relevant practice to nurture this behavior related to critical thinking (not distinguish between experts´ view and systematic literature review or importance of checking if other studies confirmed a finding of one presented research). The results are even more staggering as it shows they already know what they are doing wrong. The rationalization that "everybody does it this way" and that "we have create content fast" is not encouraging either as they are role models and trendsetters for the young generation born to the internet age. The amount of information that is consumed by young adults is gargantuan. Without their role models adopting the evidence-based approach, important decisions being made based on non-relevant or false data will become a norm affecting the decision making of all.

1. Introduction

The purpose of this research is to find out how content makers work with information sources when creating YouTube videos on scientific topics as in the pilot presented on IDIMT conference in 2019 (IDIMT, 2019). This platform counts over 2 billion of users in 2019 (Global social media ranking 2019 Statista). For several years it has been in the Top 3 of the most visited websites in the world. Additionally, YouTube is considered nowadays as a learning and educational tool for the engagement of the Web 2.0 generation (Burke and Snyder, 2008). However, the platform has no regulations regarding the validity or quality of content. According to YouTube "*Terms of Service,*" as a YouTube account holder, you should "*understand and agree that you are solely responsible for your Content and the consequences of posting or publishing it*" ("*Terms of Service - YouTube*"). Therefore, the responsibility for incorrect and misleading information is held solely by content makers. The need for the project is highly essential nowadays as YouTube becomes more popular as a new platform of knowledge (Glavas, Mathews, & Russell-Bennett, 2019).

However, unlike public media, youtube content usually does not go through the third-party editing process and the accuracy of the information communicated. Given this lack of official editing and taking into account the age of people who follow YouTubers and are therefore most influential by

them, the information provided must come from quality and valid sources. This paper aims to find out how these educational and scientific YouTubers treat sources of information in their work. Whether the communications published in scientific Youtube videos come from quality sources and are, therefore, a suitable and credible means to meet education and information needs.

Knowledge from this research could be used, for example, by marketing departments working with influencers, if they need to work with a creator with a high degree of integrity and expertise. Taking into account the results of this research, for example, secondary and primary schools could decide whether it is appropriate to include videos of scientific YouTubers as a supplement to teaching. The findings can also serve as a basis for further research into this phenomenon. Pilot of the study was presented and published in IMDT 2019 proceedings.

2. **Theoretical Background**

Social media trends in the selection of research design, data collection techniques, and analytic approaches are not well known (Snelson, 2016). All observed articles are showing the 3rd party point of view, either a viewer's opinion or an expert's assessment. Some of them focus on security point of view (Doucek, Pavlíček, et al., 2020) The point of Czech and Slovak YouTubers is covered as well. Instead of watching television, the trend of watching online videos is growing (Pereira, Fillol, and Moura, 2019). Youtube videos are discussed in several papers (Noruzi, 2017; Xiao, Wang, and Chan-Olmsted, 2018) in terms of their suitability for citation in academic papers as they are becoming increasingly popular as a source of information (Kim, Yoo-Lee & Sin, 2011). Recently, citation standards for this type of resource have even been introduced (Noruzi, 2017). Studies show that young people, in particular, consider social media as a relevant source of information (Pereira et al., 2019; Xiao et al., 2018). Some texts argue that this trend may be due to an improperly set education system (Holíková, 2018). Inefficient and rigid education does not motivate students (Pereira et al., 2019) and does not sufficiently stimulate them, especially in the current popularity of so-called infotainment.In light of these findings, a study on the interests of young people between the ages of 12 and 16 suggests paying more attention to platforms where teenagers spend much of their time - like Youtube - and explore the phenomenon of non-formal education on social networks. Involving these platforms in teaching and generating quality content could contribute to increasing teenagers' interest in learning and possibly also to increase the overall level of education in society (Aran-Ramspott, Fedele & Tarragó, 2018).School-age viewers are the largest group on Youtube (Blattberg, 2015). According to several questionnaire surveys, this target group considers YouTubers as their role models (Aran-Ramspott et al., 2018), but also sees them as trustworthy, but they are still primarily seen as a source of entertainment. Despite the characteristics of the new media, YouTubers use similar credentials to traditional media (Barden, 2018; Xiao et al., 2018). The word influencer is very often mentioned in scholarly texts - a person who can influence his viewer in some way, whether it be his views on social or political topics or his buying behavior (Aran-Ramspot et al., 2018). So if we perceive scientific YouTubers as influencers, they act as a certain authority that their audience follows and trusts. Therefore, it is important to find out if the information provided by YouTubers is true, and with what sensitivity and how the sources of information are used by scientific YouTubers. Although professional videos on Youtube are more numerous, amateur non-fiction channels enjoy higher popularity (Welbourne and Grant, 2015). The most popular researchers are life-style artists or entertainers (Aran-Ramspott et al., 2018). Therefore, the scientific side of Youtube is relatively neglected. Mostly, the research studies on downright scientific YouTubers quantitatively analyze the final output of scientific YouTubers - what formats they use to persuade, what is the link between followers and YouTuber (Welbourne and Grant, 2015; Martins-Flores and Muniz de Medeiros, 2018) or what narrative techniques are most commonly found in videos (Morcillo, Czurda, and Robertson-von Troth, 2016). These studies found that although the creators of these videos are amateurs, their

output is characterized by a high degree of sophistication, which is contrary to the general notion that Youtube content is unprofessional (Friedrich, 2017). Qualitative research is relatively neglected in this topic. Ellis-Hervey and Alston (2015), while qualitatively dealing with non-formal education videos on Youtube, are more about focusing on education about "*practical skills*" and not strictly scientific disciplines such as biology or physics. In the context of the Czech Republic and Slovakia. According to Sudzina and Pavlicek (2020) YouTubers are mostly engaged in popular media focused on youth. However, domestic scientific YouTubers are still a relatively unexplored phenomenon. Consequently, the research aims to answer the following research question: How do scientific YouTubers work with information sources? The emphasis is put on the side of content makers (YouTubers) to fill in the research gap.

3. Research Methods

The most suitable method was qualitative research, in particular in-depth, semi-structured interviews with individual YouTubers, which lasted between 40-60 minutes. The collection of information would ideally continue until the theoretical saturation is reached; however, this goal was not achieved. In-depth interviews are an ideal method for providing a large amount of complex data, which were probed to see if there are any patterns of behavior among scientific YouTubers. Semi-structured interviews provide more space for more detailed and individual information from each respondent than structured interviews (Bell, Bryman, and Harley, 2019). The aim of using this method is to obtain the most detailed and comprehensive information from the respondents. A screenplay was created with questions that focus on topics such as motivation to create, the creative process, creating credibility with the viewer, and reflecting on past work.

In order to fill the research gap, the interview method has been selected as a tool for gathering the required information in depth. Overall, the interviewing procedure has a clear purpose; it assumes preliminary planning of information gathering actions and the processing of the results obtained. The interviewing method was chosen because it has a number of advantages. First of all, it enables us to get missing information directly from the knowledge and experience of content makers. For instance, McNamara (1999) stated that interviews are particularly useful for getting the story behind a participant's experiences and that it can pursue in-depth information around the topic. Secondly, the interviewing situation, which is close to the usual form of conversation, contributes to the emergence of a comfortable atmosphere of communication and increases the sincerity of the answers. Silverman (2004) claimed that interviews are most appropriate for exploring sensitive topics. It is expected that influencers will need this setting to talk about omitting to look for reliable sources of information. And finally, the verbal communication eliminates the problem of misperception of the questions, because it is always possible to repeat and explain the question more in-depth to the person to get a clear answer. As noted by Navarro Sada & Maldonado (2007), interviewing is "*a valuable method for exploring the construction and negotiation of meanings in a natural setting.*" Furthermore, Berg (2007) argued that the value of interviewing is not only because it builds a holistic snapshot, analyses words, reports detailed views of informants, but also because it enables interviewees to "*speak in their own voice and express their own thoughts and feelings.*"

Semi-structured interviews were conducted as the purpose of it is to use discussion, conversation, as well as questions to get an insight into the research topic. According to Bernard (1988), a semi-structured interview is the best tool when there will not be any other chance to interview someone again. The choice of respondents was based on the following criteria: type of content, knowledge of Slovak or Czech, number of followers (more than 10 000).

To respect the criteria "*type of content,*" the content of YouTube videos should have been qualified as "*educational*" or "*instructional*" meaning, proving an instruction on how to do in a particular situation. Interviews were conducted personally. All interviews were recorded.

To analyze the outcome of interviews, the software MaxQDA PRO was used. The interviews were transcribed, coded, and categorized.

4. Research Sample

The basic set of respondents for the research were "*scientific and educational YouTubers,*" which are people who publish on Youtube content to educate or inform about the areas of physics, mathematics, biology, or even social sciences. The primary parameter is the assumption that these YouTube users use external sources of information to process the content of their videos.

Potential respondents were found by entering keywords and phrases such as "*science,*" "*how it works,*" "*scientific YouTuber,*" "*scientific videos*" in the Youtube search function, and the content of their video creation was then examined.

Youtubers be contacted via email and social networks. The final respondents were YouTubers who, after addressing, agreed to give an interview and to use the snowball method or referral to other suitable candidates for the interview. The final number of interviews was five.

In English, 15 respondents were interviewed for the research via Skype. They were recruited through random sampling via the search on YouTube and the table with 46 YouTubers from different countries around the world. The following keywords were used to select channels: "*scientific YouTuber,*" "*scientific channel,*" "*educational channel.*" 5 Respondents were Czech or Slovak, and they were met in person.

All respondents are English/Czech speaking YouTubers from different parts of the world. The content of their channels has an educational purpose.

5. Data Analysis and Results

MAXQDA PRO was used for coding the results. 10 interviews were coded as a pilot (Prochazka, 2019) to ensure the questions were relevant; however, they did not need to be updated. Exploration of the topic of working with resources - specifically the process of finding and assessing the quality, then what and if any activities are being developed to work credibly was tried during this period as well. During the data processing process, of course, new codes appeared, which were subsequently found in other interviews as well. Finally, 17 codes were found, which were divided into main categories - working with resources, factors influencing the creation, and influencing the viewer.

Scientific researchers generally agreed that the most critical part of video production is the research of information that precedes the creation of the script. They try to have the information they provide based on the sources and not to give their guesses and opinions as facts.

In their creation, Youtube videos are proactive and proactive in finding resources, and they go out in the field, they talk to experts on the subject. Furthermore, some experts also consult the final results with scientists.

"We have a public group, there are some people I may disagree with, who comment on it, but they are experts on the subject, or at least they can somehow oppose the subject in some way." (V., 700 odb.)

"Well, how many times do I bother my friends if they can look at me. And they tell me, "Look, I do not like that one. You said something stupid there." (M., 3200 odb.)

After conducting a series of interviews, coding was used to analyze the outcomes. In order to find out the way YouTubers work with information sources, the main categories were established - topics to be analyzed in depth. Three main categories and eight sub-categories were defined: background of a content maker, the process of working on videos, and the information source itself.

6. Findings and Discussion

The quality of resources seems to be crucial in shaping one's own opinion and the resulting message to the viewer.

"I'll open up some explanations about why, like what we have evidence for, and most of all, of course, I open up the counterparty's arguments." (K., 19,900 odb.)

"Yeah, but let us look at the methodology of those studies that say the exact opposite. Like you, they have a sample of 10 people, and those who just support my claim have a thousand people. "(M., 3200 odb.)

M. states, for example, that the quality of methodology and resources dramatically influences how he trusts the studies. The factors that YouTubers take into account when assessing the credibility and quality of resources are:

Originality of resources

"(Source) The more original, the better." (K., 19,900 odb.)

"For those articles, of course, if they have any links to other articles, I will write down what they refer to." (M., 3200, odb.)

Reputation and consistency of resources:

"The reviewed journals do not try to publish every shit, so they are a little more trustworthy than just random books." (M., 3200 odb.)

"The Guardian, it seems to be pretty consistent ..." (K., 19,900 odb.)

"when… something is released by NASA, so they have it on their official website that there must be someone stupid, that there would be something wrong from them or someone hacked and gave it wrong, but I think that it does not happen so that they publish their information that you stand behind it and I then take it from them." (F., 193 500 odb.)

Methodology:

"If you go through the methodology, as you can see how much the authors have focused on catching up like all flies ..." (M., 3200 odb.)

"So if the text is about particle physics, I suppose there was some literature review. Moreover, I suppose there will be some explanatory paragraphs in the text why it is stupid and why or not "(V., 700 odb.)

Another source of information and inspiration was the videos of other YouTubers. Interviews suggest that other YouTube creators expect a certain level of work with resources. If there are certain defects in this area, three ways of subsequent reaction can be observed.

Their own creation

"But I think they have taken this so superficially that it is no longer valid. Furthermore, I will do it better. " (M., 3200 odb)

This respondent showed the most considerable depth of work with resources and assessment of their methodologies. One possible theory is that the higher the demands of content creators on themselves, the higher they are on others.

Reporting errors to authors

"… I knew where he might have taken it, but where it came from, it indeed cannot be interpreted as he described it. … I wrote it to him, and he thanked me for it. "(M., 3200 odb.)

"… I am trying to be skeptical, even if they have it well done, but I always look for it" (F., 193 500 odb.)

Not responding

"I am pretty selfish about this. Because I do not even write on Wikipedia, nothing like that. … The articles they write, I do not know, some of the science magazines here are terrible, and like that, reporting it is difficult. "(N., 14 700 odb.)

Nevertheless, do YouTubers control themselves? Video review is either initiated by viewers in the comments. If a factual error is found, it is usually solved by a comment in the description or in the video.

"... However, I had to fix the mistake, so I wrote a comment that I highlighted in the video, I wrote it in the caption, and I did the stream ... I think the admission of those mistakes is terribly important." (K., 19 900)

For this respondent, the error was addressed on several channels.

"I will read the comments, and if there is a problem there, I find the part of the video, I fix it, I can do it, but that I would watch the video from beginning to end… I guess… I don't know if I did "(M., 3200 odb.)

"In the video about Jan Hus, I showed on the map on the wrong Husinec, because there are several… In the comments, someone corrected me, and in the end, I think I put in the description of the video that it is… "(N., 14 700 odb.)

The interaction with the viewer thus takes place not only in the comments but of course, through the content itself. As far as the other research questions about inspiring credibility are concerned, it does not seem to be crucial for YouTubers. The interviewees mainly try to show that they have worked with some reliable sources. Nevertheless, there is a strong tendency to encourage independent and critical thinking among the viewer.

"I do not want anyone to trust me just because I am supposed to be somehow trustworthy." (M., 3200 odb.)

"I tell everyone to use their brains… that is my motto because I say it, or someone I know, that does not mean it is true, and I want no one to think everything they read is credible. "(F., 193,500 subs.)

7. Conclusion

In all the found cases, quantitative studies have been carried out on the topic of scientific YouTubers, however, this research also contains some points that could be related to their results. Appropriateness of using Youtube videos as a source of information for work and education at schools could be controversial. In some cases YouTubers use scientific articles and/or expert advice for their videos.

It seems, however, that expert advice is a more common practice. This finding supports a survey on Brazilian YouTubers (Martins-Flores and Muniz De Medeiros, 2018).

The principal limitation of this research is that generalization of any kind cannot be made.

The direction of further research could be more focused on the interaction and the real impact of the videos on the viewer, be it critical and independent thinking, as scientific YouTubers have mentioned several times that there are values they would like to convey to their viewers.

Subsequently, one could examine the influence of the chosen YouTubers' formats on the memorization of information or the effectiveness of learning from these types of videos.

The participative observation should undergo as follow-up research. Comments linked to the videos could be coded and categorized. The main strength of the observation method is that it gives direct facts on whether some conditions are fulfilled or not. In addition, due to the stated above limitations, a study in depth of one scientific field could show the similarities and differences in using valid and reliable sources of information and testing possible correlations. The wider study that would confirm these preliminary findings should follow (Prochazka, 2019)

8. References

Alston, G. D., & Ellis-Hervey, N. (2015). Exploring the nonformal adult educator in twentyfirst century contexts using qualitative video data analysis techniques. Learning, Media and Technology. https://doi.org/10.1080/17439884.2014.968168

Aran-Ramspott, S., Fedele, M., & Tarragó, A. (2018). Youtubers' social functions and their influence on pre-adolescence. Comunicar, 26(57), 71-79. doi:http://dx.doi.org/10.3916/C57-2018-07

Barden, J. (2018, May 1). Where's the fair use? Participatory culture, creativity, and copyright on youtube. (Disertační práce). University of Wisconsin-Milwaukee. https://dc.uwm.edu/etd/1744/

Bell, E., Bryman, A., & Harley, B. (2019). Business research methods. Oxford, Anglie: Oxford University Press.

Berg, B. (2007) An Introduction to Content Analysis. In: Berg, B.L., Ed., Qualitative Research Methods for the Social Sciences, Allyn and Bacon, Boston, 238-267.

Blattberg, E. (2015, April 24). The demographics of YouTube, in 5 charts. [vid. 2019-03-29]. https://digiday.com/media/demographics-youtube-5-charts/

Burke, Sloane C.; Snyder, Shonna L.(2008) YouTube: An Innovative Learning Resource for College Health Education Courses. International Electronic Journal of Health Education, v11, 39-46

Doucek, P., Pavlíček, L., Sedláček, J., & Nedomová, L. (2020). Adaptation of Password Strength Estimators to a Non-English Environment – the Czech Experience. Computers & Security, 101757. doi:10.1016/j.cose.2020.101757

Kim, K.-S., Yoo-Lee, E., & Sin, S.-C. J. (2011, January 11). Social media as information source: Undergraduates use and evaluation behavior. Proceedings of the American Society for Information Science and Technology, 48(1), 1-3. doi:10.1002/meet.2011.14504801283

Glavas, C., Mathews, S., & Russell-Bennett, R. (2019). Knowledge acquisition via internet-enabled platforms: Examining incrementally and non-incrementally internationalizing SMEs. International Marketing Review, 36(1), 74–107. https://doi.org/10.1108/IMR-02-2017-0041

Global social media ranking 2020 Statistic. (n.d.). Retrieved March 28, 2020, from Statista website: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

H. Russell Bernard: Research Methods in Cultural Anthropology 1988, Beverly Hills: Sage. 520 pages. John Van Maanen: Tales of the Field: On Writing Ethnography 1988, Chicago: Chicago University Press. 173 pages. Organization Studies, 11(3), 464–466.

Martins-Flores, N., & Muniz De Medeiros, P. (2018 December 31). Science on YouTube: Legitimation Strategies of Brazilian Science YouTubers. Revue Française Des Sciences De L'information Et De La Communication,(15). doi:10.4000/rfsic.4782

Morcillo, J. M., Czurda, K., & Robertson-von Trotha, C. Y. (2016 May 25). Typologies of the popular science web video. Journal of Science Communication, 15(4). doi:10.22323/2.15040202

Navarro Sada, A., & Maldonado, A. (2007). Research Methods in Education. Sixth Edition - by Louis Cohen, Lawrence Manion and Keith Morrison. British Journal of Educational Studies, 55(4), 469–470. https://doi.org/10.1111/j.1467-8527.2007.00388_4.x

Noruzi, A. (2017, June). YouTube in scientific research: A bibliometric analysis. Webology, 14(1), 1-7. http://www.webology.org/2017/v14n1/editorial23.pdf

Pereira, S., Fillol, J., & Moura, P. F. R. de. (2019). Young people learning from digital media outside of school: The informal meets the formal. Comunicar, 27(58), 41-50. doi:http://dx.doi.org/10.3916/C58-2019-04

Prochazka, D. A. (2019). WHY DON'T YOU CHECK YOUR SOURCES? THE NONEVIDENCE-BASED APPROACH OF YOUTUBE INFLUENCERS . In: IDIMT-2019, Innovation and Transformation in a Digital World (pp. 251–258). Weinheim, Germany: Beltz Verlag.

Silverman, D. (2004). Interpreting qualitative data: methods for analysing talk, text and interaction (2. ed., reprinted). London: Sage Publ.

Snelson, C. L. (2016). Qualitative and Mixed Methods Social Media Research: A Review of the Literature. International Journal of Qualitative Methods, 15(1), 160940691562457. https://doi.org/10.1177/1609406915624574

Sudzina, F.; Pavlicek, A. Virtual Offenses: Role of Demographic Factors and Personality Traits. Information 2020, 11, 188.

Terms of Service - YouTube. (n.d.). Retrieved April 8, 2020, https://www.youtube.com/t/terms?gl=CZ

Welbourne, D. J., & Grant, W. J. (2015, February 19). Science communication on YouTube: Factors that affect channel and video popularity. Public Understanding of Science,25(6), 706-718. doi:10.1177/0963662515572068

Xiao, M., Wang, R., & Chan-Olmsted, S. (2018, July 30). Factors affecting YouTube influencer marketing credibility: A heuristic-systematic model. Journal of Media Business Studies,15(3), 188-213. doi:10.1080/16522354.2018.1501146

BIG DATA ETHICS AND SPECIFIC DIFFERENCES FROM GENERAL DATA ETHICS

Richard Novák, Antonín Pavlíček

Faculty of Informatics and Statistics, University of Economics, Prague
Richard.Novak@vse.cz, Antonin.Pavlicek@vse.cz

Keywords

big data, data ethics, issues, macro-ethics, regulation, stakeholders, use cases

Abstract

Big Data ethics is a subset of data science and applied ethics; however, there are some aspects that make this topic very unique and that we discuss in the paper. Firstly, we start with a literature overview that covers several areas that are relevant to Big Data ethics such as computer ethics, professional ethics, information ethics and data ethics. Secondly, we dive deeper into data ethics viewed as macro-ethics consisting of the ethics of data, algorithms and practices. Finally, we name and describe the Big Data ethics specifics such as the specific role of stakeholder groups, opportunities arising from the new use cases, challenges (issues) and conflicts stemming from Big Data and also the new demand for a regulatory framework.

1. Introduction

This paper will discuss Big Data in the context of data science[5] and applied ethics. At the heart of the paper is the following quote from (Sokol, 2016) and (Boyd & Crawford, 2012, p. 671)

"Just because it is possible does not make it ethical."

We can observe that people's insight into complex problems and their attitudes towards life is currently driven by advanced technologies such as Big Data and Artificial Intelligence among others. It means *"Change the instruments, and you will change the entire social theory that goes with them "*. (Latour, 2009, p.155).

The free use of technology such as internet, mobile phones, navigations, IoT and some others generating Big Data masks its danger to society as some respected authors such us Floridi (2016), Sokol (2016) and Boyd & Crawford (2012) warn us.

Thus, university researchers are formulating new hypotheses about the possible shifts, divides, manipulation and inequalities in society driven by new technologies and that is leading to the current trend to view Big Data as socio-technological phenomenon, (Boyd & Crawford, 2012).

The thesis is summarizing the previous research on Big Data ethics considering the relevant literature and providing a new perspective that is adding a few specifics to Big Data ethics.

---

[5] Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured, similar to data mining. (Wikipedia, 2020)

1.1. Elementary Definitions

There are a few good definitions of Big Data. The most common are from consulting companies like this one from Gartner (focusing on big volume, variety and velocity), however we found as more relevant definition of Big Data for this paper that comes from the Microsoft Research Center:

"We define Big Data as a cultural, technological, and scholarly phenomenon that rests on the interplay of:

1. *Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets*

2. *Analysis: drawing on large data sets to identify patterns in order to make*

economic, social, technical, and legal claims

3. *Mythology: the widespread belief that large data sets offer a higher form of*

intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy". (Boyd & Crawford, 2012, p. 663)

We should also describe here, in the definition part, the terms of Ethics to clarify the broader context of the paper that we will dive deeper into throughout the following chapters.

The term ethics derives from Ancient Greek and was firstly used by Aristotle to name a field of study developed by his predecessors Socrates and Plato.

"Philosophical ethics is the attempt to offer a rational response to the question of how humans should best live." (Wikipedia, Aristotle, 2020)

There are many areas of ethics such as meta-ethics or normative ethics, although for this paper applied ethics is the most relevant because it can further be adapted to more detailed fields such as bioethics, business ethics or Big Data ethics.

"Applied ethics is a discipline of philosophy that attempts to apply ethical theory to real-life situations." (Ethical World, 2018)

2. Big Data Ethics Overview

2.1. General Overview

Applied ethics covers several areas that are relevant to Big Data such as **computer ethics**, **professional ethics**, **information ethics** and **data ethics**.

Before we discuss, the above mentioned, relevant areas of applied ethics, we will do a short overview of more general terms such as digital ethics, cyber ethics and business ethics.

Digital ethics is an umbrella term covering all issues raising from the conflict of digital technologies and ethics. Robert Capurro defines digital ethics as:

"Digital ethics or information ethics in a broader sense deals with the impact of digital Information and Communication Technologies (ICT) on our societies and the environment at large." (Capurro, 2009).

Cyber ethics covers, in our view, the behavior in a broader area of virtual cyber space of society that is created by ICT. For a definition of cyber ethics, we can use the following:

"Cyber ethics is a set of moral choices individuals make when using internet-capable technologies and digital media." (Chen, 2012).

For a more detailed description of cyber ethics, see Tomas Sigmund´s (2013) article: *Ethics in the Cyberspace*.

Business ethics is, for us, an even more general term covering classical philosophy challenged by business environments and moral problems arising from different issues such as conflicts of interest, social contracts and stakeholder groups.

This paper is focused on Big Data ethics thus our approach goes from general business ethics to specific professional ethics of data science related to impacted people. Furthermore, our approach goes from general data science to the specifics of Big Data. For a view on the possible intersection of data science with cyber space, which is not our direct focus, we recommend the book by Richard A. Spinello, (2010), *Cyberethics: Morality and Law in Cyberspace*.

In the following text, we discuss the relevant Big Data ethics predecessors such as computer ethics, professional ethics, information ethics and finally also data ethics.

2.2. From Computer to Information Ethics

Computer ethics has been evolving since the invention of computers in the 20th century after the world wars. A deeper look at the moral problems related to information technologies started in the 40s by the work of MIT professor Norbert Wiener that introduced the term "Cybernethics" in his eponymous book (Wiener, 1948). It was followed by his other books. For example, in *The Human Use of Human Beings*, Wiener explored some likely effects of information technology upon key human values like life, health, happiness, abilities, knowledge, freedom, security, and opportunities (Bynum, 2015).

Wiener's findings were followed almost three decades later by many authors using the term computer ethics such as Walter Maner or James Moor. Maner stated in his article *Starter Kit in Computer Ethics*, (Maner, 1980) that, *"Wholly new ethics problems would not have existed if computers had not been invented."* And Moor in his article *What is Computer Ethics?* defined this area with the following:

"In my view, computer ethics is the analysis of the nature and social impact of computer technology and the corresponding formulation and justification of policies for the ethical use of such technology, (Moor, 1985).

In a nutshell, computer ethics is focused mainly on new technology innovation, meaning new tools and machines: *"When computers interact with society they are causing new moral problems."* (Friedman and Nissenbaum, 1996). From the point of the involved stakeholders, computer ethics focuses on the relationship of public users and computer professionals and their professional ethics covering: *"Standards of good practice and codes of conduct for computing professionals."* (Gotterbarn, 1991).

On the other hand, the terms of information ethics and data ethics that have been used for the last decade by respected authors like R.Cappuro, L.Floridi and M.Taddeo focus on either more a philosophical approach (Capurro, 2006) or on the content at different levels of abstraction (Floridi, 2008). Content was originally considered different entities such as knowledge, information and data, although the latest developments have moved the focus on raw data that is the new target of our moral actions described as:

"The shift from information ethics to data ethics is probably more semantic than conceptual, but it does highlight the need to concentrate on what is being handled as the true invariant of our concerns. This is why labels such as 'robo-ethics' or 'machine ethics' miss the point, anachronistically stepping back to a time when 'computer ethics' seemed to provide the right perspective." (Floridi & Taddeo, 2016).

2.3. Information and Data Ethics

The foundation of modern information ethics was laid down at the end of the 20th century by Rafael Capurro and Luciano Floridi. Nowadays, this theme is further developed not by individuals but mainly by broader working groups concentrated at university labs with a special focus on data ethics such as Oxford University (Digital Ethics Lab, Floridi´s workplace), Utrecht University (Ethical Institute) and Vienna University (The Privacy and Sustainable Computing Lab).

It seems that the more ICT specific approach of Floridi prevails in the academic and also the business world, although we would like to mention also the more general philosophical approach of R. Capurro that provides broader philosophical perspectives of ontology and metaphysics.

Although Capurro and Floridi do not understand each other in their approaches they both are going the same direction. We will further discuss Floridi´s approach because it is currently the leading view in data science, and at the same time, it is easy to understand, especially for the relevant area of Big Data ethics.

In his article *Information Ethics, Its Nature and Scope* (2006), Floridi suggested the unified approach towards information ethics that he calls macroethics. It consists of three arrows with information as a source, information as a product and information as a target. He also introduced the idea of a moral agent that can generate the information as a product and affect the information environment as a target (Floridi, 2006); (Sigmund, 2013).

The most recent definition of data ethics was done in 2016 by two Oxford academics, Luciano Floridi and Mariarosaria Taddeo, approaching the topic again as macroethics distinguishing the **ethics of data**, **algorithms** and **practices**.

This respected definition of data ethics in Level of Abstraction of data (LoA$_D$) was done in their article *What is data ethics?* (2016) as:

"In the light of this change of LoA, data ethics can be defined as the branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values). This means that the ethical challenges posed by data science can be mapped within the conceptual space delineated by three axes of research: the ethics of data, the ethics of algorithms and the ethics of practices." (Floridi & Taddeo, 2016).

3. Big Data Specifics

We see progress in the relevant data ethics areas described above in the latest work of Floridi and Taddeo; however, in our opinion, there are some Big Data specifics that need to be discussed in more detail.

By these specifics of Big Data, we name the following that we will discuss further:

- Specific role of stakeholder groups (organizations, users, states).

- Use cases of Big Data, (showing mainly positive benefits)

- Conflicts and issues stemming from Big Data (Digital Divide and other dilemmas)

- Demand for regulatory framework (very different approach in EU, USA and Asia)

3.1. The specific role of stakeholder groups

It means that there is increasing information asymmetry between the individual users that can be considered data poor[6] and big corporations that collect data about the individual users and can be considered data rich. To be data rich means to have data insight into many areas as all of society is getting "datafied"[7] and this data insight leads to many advantages in the form of new business opportunities. A strong role from nation states is expected to regulate this information asymmetry and balance the equal opportunities and basic human rights. However, the power of some corporations derived from their turnover is exceeding; in the cases of the biggest corporations such as Google, Facebook, Microsoft, Apple or Alibaba; the state budgets of many nation states. See the figures at (https://www.statista.com) to compare the actual state budget of European countries and the turnover of leading global corporations that collect data about their users.

Besides the role of data poor users and data rich states and organizations there is important role of the so called "power users" or in another words "IT specialist" (Data scientists, Data engineers). They are designing the complex Big Data systems when working for states and organizations. That is why, there is very high importance of the self-regulatory approach applied by IT specialist in the area of Business and Professional ethics highlighted in previous chapters, see e.g. Spinello, (2010).

3.2. Use cases

Use cases demonstrate the new possibilities of Big Data technologies that are followed by generated issues (risks and issues), however, that use cases stress mainly the positives sides of Big Data.

Despite the relative youth of the use cases field – the first studies started less than 20 years ago – it is becoming a powerful tool for understanding of patterns and processes in human and also machines behaviors. It allows to identify algorithms and patterns that often create competitive advantage and provides unexpected views and facts hidden in data lakes with high entropy.

Among the benefits of use cases, we can briefly name the following:

Insight into complex problems, new business opportunities, increased security, decreased costs thanks to better analytics, speeding up complex decisions or more targeted and personalized offers, and more.

There are many articles describing the use cases from different industries such as: telecommunication (Pavlicek, Doucek, Novak, Strizova, 2017), finance (Fang & Zhang, 2016). among others.

3.3. Issues and conflicts stemming from Big Data

We can see that Big Data use cases and human values[8] are powers that are inherently in a conflict. Few basic examples of the possible conflicts between human values and Big Data benefits named in previous chapter are shown below:

- Privacy   vs   New business opportunities

---

6   The term information asymmetry and terms data poor and data rich are used by many authors, for example, Boyd & Crawford, (2012)

7   The term "datafication" means that we create digital data about almost every existing subject and was firstly defined in 2013 by K.N.Cukier and V. Mayer-Schoenberger.

8 The human values and rights are a special area of interest in sociology and moral philosophy. For the purpose of this article we will follow the human values as they are defined in declaration of EU and stated in e.g., (European parliament, 2000)

- Privacy   vs   Insight into complex problems
- Equality   vs   Digital divide …etc.

The scope of this thesis does not allow for the discussion of the possible conflicts in detail. There are special papers discussing the conflicts stemming from the implementation of Big Data and other related technologies. For more details in this area see, e.g. (Whittlestone at el, 2019).

Big Data can be considered a new technology causing similar issues (digital divide…etc.) as the internet and other technological innovations that are already well described in literature, see e.g. Norris, (2001), Dijk, (2006) among others.

Thus, it is a good question: what is really new in the phenomenon of Big Data? We can start the insight into Big Data issues with the reputable work of Boyd and Crawford named: CRITICAL QUESTIONS FOR BIG DATA, which can be summarized in the so called Six Provocations of Big Data listed below:

- "Big Data changes the definition of knowledge

- Claims to objectivity and accuracy are misleading

- Bigger data is not always better data

- Taken out of context, Big Data loses its meaning

- Just because its accessible does not make it ethical

- Limited access to Big Data creates a new digital divide." (Boyd and Crawford, 2012).

Although the six provocations create a good foundation for the list of Big Data issues, we recognize that more of them exist. We researched comprehensively Big Data issues and we discover a extended list of twelve issues described, e.g. in a work of Novak (2020) that aggregated the finding of many relevant authors.

Here is a comprehensive list of twelve Bid Data issues: *"Privacy Intrusion, New Barriers, Business Advantage, Power of All data, New Big Brother effect, Missing Transparency, Confusion, Social Pressure, Belief in Legislation, End of Theory, Data Religion and Unawareness of our Data.",* (Novak, 2020).

It would be interesting to discuss each of the one from twelve issues above separately; however, the scope of our paper does not allow it, thus we recommend for more details to approach directly the original work of Novak (2020).

3.4. Demand for regulative framework

The demand for regulative framework were concluded in the EU with the General Data Protection Regulation (GDPR) that became enforceable beginning 25 May 2018; however, there are also other contexts and means than just GDPR that can regulate the use of Big Data in society.

The American lawyer and professor of law at Harvard university, Lawrence Lessig, described several causes of possible regulations of cyberspace such as the **market, legislation, social norms** and **architecture** in his book *Code and Other Laws of Cyberspace,* (1999).

The deep dive into regulation framework of Big Data and new proposals improving the current inequalities in society is definitely needed. Authors of this article are working on a special article dedicated to this topic; however, it is definitely beyond the limit of this article. We recommend the work of Lessing (1999) named above for the interested and for this moment we consider the overview provided above and available in, e.g. Lessig´s books as the pragmatic compendium.

4. Summary of Big Data Ethics Overview

We agree with the latest development of data ethics done in 2016 in the work of Floridi and Taddeo that is for us a logical evolutionary step from other authors such as Wiener, Maner, Moor, Gotterbarn that focused on the relation of men and machines (computer and professional ethics). We also believe that Capurro's theoretical work in the area of information ethics with great focus on the broader concept of information ethics and digital ontology contributed to the current founding landscape of data ethics done by Floridi and Taddeo.

Their approach is based on different levels of abstractions (LoA). This view of data ethics changes the focus from information to raw data. It offers a flexible and complex approach to the topic. Their ideas of dealing with data ethics as macro-ethics consisting of the **ethics of data**, **algorithms** and **practices** is a solid attempt to solve the complexity of this area avoiding the narrow and ad-hoc approach.

In spite of agreeing with Floridi and Taddeo's work in data ethics, we find Big Data ethics as rather different from general data ethics. Describing these specifics and their implication is, as we believe, our contributions done in this article.

We named the Big Data ethics specifics in the chapters above, such as the specific role of stakeholder groups, opportunities (benefits) arising from the new use cases, challenges (issues) and conflicts stemming from Big Data and also the new demand for a regulatory framework.

5. Acknowledgement

6. References

Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA. In: Information, Communication & Society. (pp. 662-679). Volume 15: Issue 5.

Bynum T. (2015). Computer and information ethics. In The Stanford encyclopedia of philosophy (ed. EN Zalta), Winter 2015.

Capurro, R. (2006). Towards an ontological foundation of information ethics. Ethics and information technology, 8(4), 175-186.

Cukier, Kenneth; Mayer-Schoenberger, Viktor (2013). "The Rise of Big Data". Foreign Affairs (May/June): 28-40. Retrieved 24 January 2014.

Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. Poetics, 34(4-5), 221-235.

European Parliament. (2000). Charter of fundamental rights of the European Union. Office for Official Publications of the European Communities.

Fang, B., & Zhang, P. (2016). Big data in finance. In Big data concepts, theories, and applications (pp. 391-412). Springer, Cham.

Floridi, L. (2006). Information ethics, its nature and scope, ACM SIGCAS, 36 (3)

Floridi, L., & Taddeo, M. (2016). What is data ethics?

Friedman, B. and H. Nissenbaum (1996), "Bias in Computer Systems," ACM Transactions on Information Systems, 14(3): 330–347.

Gotterbarn, D. (1991, May). Ethical considerations in software engineering. In Proceedings of the 13th international conference on Software engineering (pp. 266-274).

Chen, J. (2012), Cyberethics. University of British Columbia. January 2012. http://etec.ctlt.ubc.ca/510wiki/Cyberethics.

Latour, B. (2009). 'Tarde's idea of quantification', in The Social After Gabriel Tarde: Debates and Assessments, ed M. Candea, London: Routledge, pp. 145-162.

Lessig. L., (1999). Code and Other Laws of Cyberspace. Basic Books, Inc., NY, USA.

Maner, W. (1980). Starter kit in computer ethics. Hyde Park, NY: Helvetia Press and the National Information and Resource Center for Teaching Philosophy.

Moor, J. H. (1985). What is computer ethics?. Metaphilosophy, 16(4), 266-275.

Norris, P. (2001). Digital Divide: Civic Engagement, Information Poverty and the Internet in Democratic Societies. New York: Cambridge University Press.

NOVÁK, R. (2014-2020). Big Data and Ethics (Doctoral dissertation, Vysoká škola ekonomická v Praze).

Pavlicek, A., Doucek, P., Novák, R., & Strizova, V. (2017). Big data analytics–geolocation from the perspective of mobile network operator. In International Conference on Research and Practical Issues of Enterprise Information Systems (pp. 119-131). Springer, Cham.Ethical world. (2018).

Sigmund, T. (2013). Ethics in the Cyberspace. IDIMT-2013 Information Technology Human Values, Innovation and Economy, 42, 269-279.

Sokol, J. (2016). Ethic, Life and Institutions, Praha: Vyšehrad

Spinello, R. (2010). Cyberethics: Morality and law in cyberspace. Jones & Bartlett Learning.

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019, January). The role and limits of principles in AI ethics: towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 195-200). ACM.

Wiener, N. (1948). Cybernetics. Scientific American, 179(5), 14-19.

PARENTAL ACTIVITY ON SOCIAL NETWORKS

Jana Syrovátková, Richard Novák

Faculty of Informatics and Statistics, University of Economics, Prague
Jana.Syrovatkova@vse.cz, Richard.Novak@vse.cz

Keywords

parenthood, social network, parental activity, child age

Abstract

Parents of young children are frequent and grateful target of internet marketing. For accurate targeting, it is important to know when users are the most active on the Internet, given the age of the children, and when, on the contrary, they are fully engaged in childcare. For the analysis, we chose the "blue horse" social network focused on parents and children. We further compared "term groups" - ie. discussion for parents with the date of birth of a child in a particular month.

1. Introduction

1.1. State of the art

Internet marketing studies many areas of user behavior. One such commonly addressed topic is e.g. the optimal time for publishing new content. In (Pavlíček, Doucek, Novák, 2018) the authors address the whole content of Facebook posts, while in (Sudzina, Pavlíček 2017) the authors discuss how gender impacts orders from various deal sites. There are studied social network from various perspectives of use in companies (Pavlíček, 2010, 2014).

This article examines behavior of a specific user group – parents of young children. It discusses how / whether the age of children impacts the internet activity of the parents, namely their behavior on social networks frequented by pregnant women and young children parents. To the author´s knowledge, this topic has not been covered by an existing literature.

1.2. Blue horse

Blue Horse is a social network / discussion forum created on June 30, 2006. The main topics are children and pregnancy. It brings together mostly users and discussions and questions about children and motherhood. A large part of the Blue Horse are discussions of women trying to conceive, undergoing artificial insemination and sharing their pregnancies. (Blue Horse, 2018)

The blue horse contains everything that ordinary social networks contain. Each registered user has their own profile on the Blue Horse. He can have his profile photo in it and reveal a little about himself. The user can publish short news, albums and articles on their profile. At the same time, they will appear on the walls to others who liked the user. For social networks and therefore also the Blue Horse, private messages between users are a matter of course (Blue Horse help, 2020).

Several different formats are used for discussions. For simple questions, there are thematically structured discussions, referred to as forums. You can watch current threads on the walls of other users. Furthermore, they can be grouped according to many different criteria. It can be the same

hobby, experience, problem, residence, date of birth. Each user can create their own group. Unlike walls or discussions, group threads are only available to group members. Groups can be public - anyone can join them – or private – there is an invitation necessary (Bluehorse help, 2020).

A very interesting part of the Blue Horse is the bazaar, where you can sell and buy everything related to children and families. For fun, there are many competitions for interesting prizes and the opportunity to test products designed for pregnant women or young children.

Many users are bothered by specific issues, connected with the childcare and usually the domain of childcare professionals. Many answers to questions can be found on the wiki. This is an encyclopedia of what bothers and interests women, mothers and their children. The articles are more professional and try to maintain an independent view / attitude as is common for wikis (Pavlicek, Rosicky, 2008). If they can't find the answer on the wiki, Blue Horse still contains numerous professional counseling chat rooms. Here it is possible to ask the question directly to experts – doctors, lactation consultants, dentists, midwives, gynecologists and many others.

1.3. **Term groups**

Pregnancy in the first trimester is very risky. Many pregnancies end in the first trimester with miscarriage or the fetus stops developing and the pregnancy must be terminated. Many mothers are so reluctant to say this news at work, in their families. Nevertheless, they long to share fears and gradually joys with someone. Discussions of mothers with a similar date of childbirth often search for anonymous internet environments.

As mentioned above – groups are intended for women connected by a common due date. Even outside the Blue Horse, discussions connecting women with the same month of childbirth are often called term groups. Groups are thus typically marked with the month and year of the planned due date.

The blue horse will make this group available to the user automatically after entering his due date. Some users view the group a month or two before their due date. A frequent reason is the term at the turn of the months.

We will therefore continue to use groups exclusively for this text. There are always individual threads in groups. Responses to threads are listed directly below the relevant thread. It is easy to see the author's nickname, date, thread, number of likes and number of comments on the thread.

2. **Goals**

The aim was to obtain data from multiple term groups and compare these. The main goal was to find out how active the individual groups are during their functioning. When are they most active and when are they less active? Whether any trends can be traced in all groups the same. The long-term activity of the group is also an interesting figure (Pavlíček, Doucek, 2017).

We also focused on individual nicknames. We examined how many users actually discuss in multiple groups, how many users discuss in the long run. An interesting parameter was the number of contributions of individual users.

3. **Methods**

The first step was to download the html codes of all pages of the relevant discussion. Initially, we downloaded it manually. For simplicity and to determine the error rate, we then generated an html page with links to the entire discussion and used the DownThemAll tool. This way we got a lot of

html files (Černý, Potančok, Molnár, 2019). Unlike Facebook, data can also be easily downloaded from closed groups (Pavlíček, Doucek, Yablotschnikov, 2019).

Visual Basic for Excel was used to convert html files to the database. We create a short program that opened html in Excel, copied data from it and then identified the date of the thread, nickname, number of comments and the number of likes, which it wrote into the database.

Pivot tables and contingency graphs were primarily used for the analysis of the obtained data (Mládková, Procházka, 2015).

4. Analysis

Four term groups were used for the analyzes - February, March and April 2019 and September 2018. Three groups were used side by side for possible analysis of users discussing in several groups. The fourth group was for removing the influence of the seasons.

In each of the groups, the first contributions are 8 to 9 months before the respective deadline. Ie. it can be seen that women typically start discussing at an early stage of pregnancy.

There are different numbers of threads in the groups. It is understandable that a longer functioning group has more contributions. The March 2019 group has a relatively high number of entries, which can be explained by competitions in the group. It is interesting that the number of comments and likes of individual threads is similar in all groups. At the same time, each user can equally give their own thread and comment.

Term group	Number of threads	Number of like	Number of comments	Like to threads	Comments to threads
9.18	6 011	53 076	53 491	8,83	8,90
2.19	2 950	23 083	26 607	7,82	9,02
3.19	4 559	33 594	38 861	7,37	8,52
4.19	2 175	15 284	18 066	7,03	8,31
Total	**15 695**	**125 037**	**137 025**	**7,97**	**8,73**

Table 1 - Global numbers for all groups

4.1. Analysis of threads

The main purpose of the analysis was to examine the activity of users in the relevant months before, around and after childbirth. Graphs of the number of threads, the number of comments and the number of likes in the respective months are used for this purpose.

In the Czech Republic, a woman takes maternity leave 6-8 weeks before the planned due date. The graphs show a sharp increase in user activity about 2 months before that, ie. just in connection with taking maternity leave. Maximum activity is at the moment of the birth of a child. This is followed by a gradual decline. The most significant decrease is in the number of likes. For contributions and comments, the decline is first sharp and then slow until about 10-11 months of age. In the child's year there is a slight increase, but then a decrease again. Interesting is the increase around the age of 15 months in the group of children born in September 2018. This would be interesting to explore in future research.
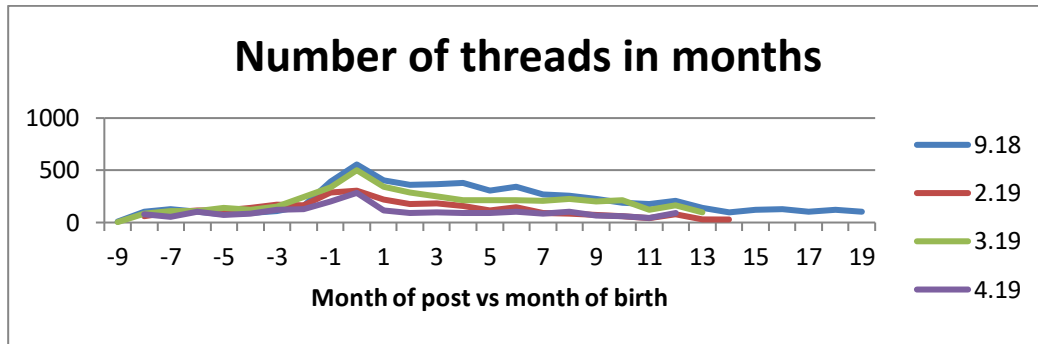
Chart 1 - Number of threads in months
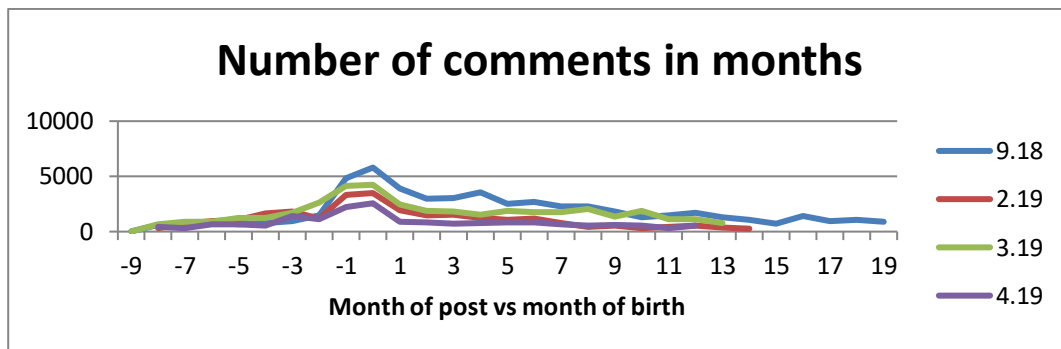
Chart 2 - Number of comments in months
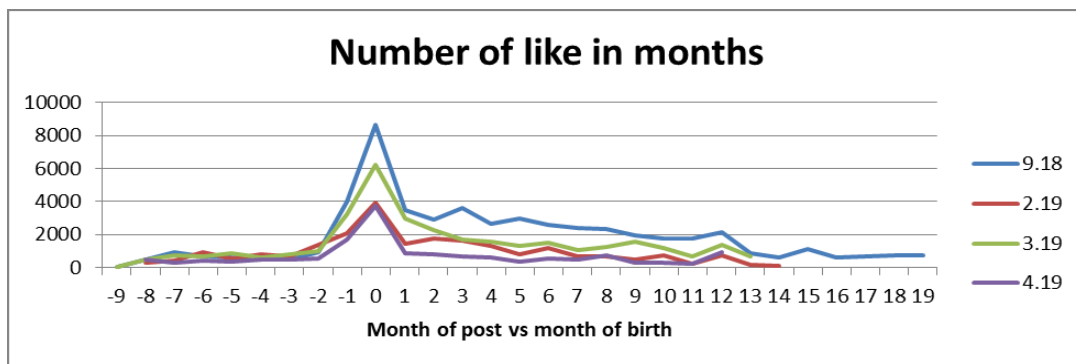
Chart 3 - Number of like in months

4.2. Analysis of users

There are 840 of individual users in these 4 groups. There are 229 users in the March 2019 group and only 35 of them are in more groups - in February 2019 or in April 2019 group. This shows that most users discuss only in one group connected with their expected due date.

Another interesting question is whether the users discuss regularly or whether most users are in the group for a rather short time. The chart below shows that most users have only 1-5 threads during the group. Another significant group is 6-10 threads. Only less than half of the users discuss in the group for a long time. On the other hand, from the next graph it can be seen that the curves of the number of users who have at least one contribution in a month decrease very slowly. Thus, it can be said that many users are not in the group from the beginning, but rather come randomly during the duration of the group.

Term group	Number of nicknames
9.18	279
2.19	176
3.19	229
4.19	192
more groups	35
Total	**840**

Table 2 - Number of active nicknames

Threads	9.18	2.19	3.19	4.19
1-5	124	88	105	97
6-10	34	20	34	36
more than 10	121	68	90	59

Table 3 - Number of threads of individual users
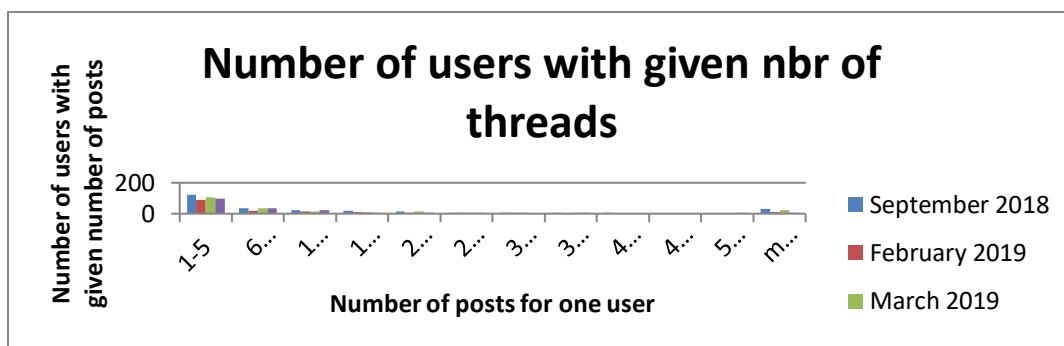
Chart 4 - Number of users with given nbr of threads
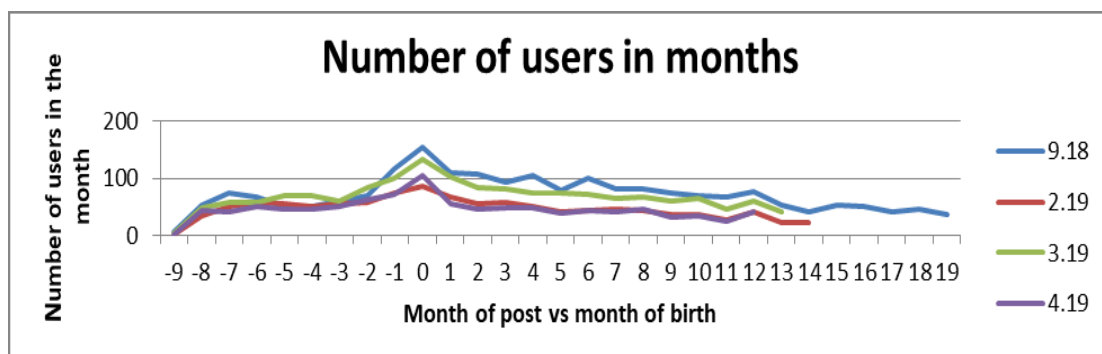
Chart 5 - Number of users in months

5. Discussion and conclusion

It is interesting to note the sharp increase in the number of contributions as well as the sharp decrease later. There is an understandable increase in the time of the children's birthday, but even that only balances the numbers around 9 months of age of the child. There is a really significant increase at a time when women are on maternity leave. It can be said that the big decline after that is at the moment

when the children start to climb and walk. Ie. from the point of view of targeting possible marketing, it makes sense to focus on women after taking parental leave and women with small babies.

Within the analysis of users, it is interesting that most discussants only look into the group or are passive in it and do not make their own contributions. Due to the slow decrease in the number of different users discussing in a month, this can perhaps be explained by the fact that users go for advice with various problems with children. And various problems come in time regardless of the age of the children.

6. References

ČERNÝ, J., POTANČOK, M., MOLNÁR, Z. (2019) Using open data and Google search data for competitive intelligence analysis. Journal of Intelligence Studies in Business JISIB [online]. 9/2, pp. 72–81. ISSN 2001-015X. available at: https://ojs.hh.se/index.php/JISIB/article/download/470/216.

MLÁDKOVÁ, L., PROCHÁZKA, D. A. (2015) Leading Change in Management Education in Department of Management. In: ROUCO, José Carlos Dias (ed.). Proceedings of the 11th European Conference on Management Leadership and Governance. Lisabon, 12.11.2015 – 13.11.2015. Reading: Academic Conferences and Publishing Int. Ltd., pp. 248–253. ISBN 978-1-910810-76-7. ISSN 2048-9021, WOS:000380511600031

BLUE HORSE (2018). Jak omylem vytvořit web pro ženy s milionem uživatelů měsíčně – Modrý koník marketérům. https://www.modrykonik-marketerum.cz/blog/jak-omylem-vytvorit-web-pro-zeny-s-milionem-uzivatelu-mesicne/

BLUE HORSE HELP (2020). Potřebuji se na Modrém koníku zorientovat. Modrý Koník .cz. https://www.modrykonik.cz/faq/potrebuji-se-na-modrem-koniku-zorientovat/

PAVLICEK, A. (2010). Human Resources in ICT- Information and Technological Literacy as the Qualification Prerequisites. In P. Doucek, G. Chroust, & V. Oskrdal (Eds.), IDIMT-2010: INFORMATION TECHNOLOGY - HUMAN VALUES, INNOVATION AND ECONOMY (Vol. 32, pp. 153–159). Universitatsverlag Rudolf Trauner., ISBN: 978-3-85499-760-3 WOS:000288345500016

PAVLICEK, A. (2014). New and Social Media in Workplace. In IDIMT-2014: NETWORKING SOCIETIES - COOPERATION AND CONFLICT (pp. 135–144). Trauner Verlag. ISBN 978-3-99033-340-2. WOS:000346148300015

PAVLICEK, A., ROSICKY, A. (2008). Wiki System as a Knowledge Management Tool. In 16th Conference on Interdisciplinary Information Management Talks (Vol. 25, pp. 403–412). Universitatsverlag Rudolf Trauner. WOS:000265436800030

PAVLÍČEK, A., DOUCEK, P., NOVÁK, R. (2018). Facebook Posts Engagement Analysis – Case Study of the Leading e-Shop in the Czech Republic. In: Research and Practical Issues of Enterprise Information Systems. Poznań, Cham: Springer, pp. 27–39. ISBN 978-3-319-99039-2. DOI: 10.1007/978-3-319-99040-8_3.

PAVLÍČEK, A., DOUCEK, P., YABLOTSCHNIKOV, S. (2019) Regional Differences in Facebook Privacy Settings and Behaviour. In: Conference of Open Innovations Association (FRUCT). Moscow, RUS : IEEE, 2019, pp. 314–320. ISSN 2305-7254. DOI: 10.23919/FRUCT.2019.8711954

PAVLÍČEK, A., DOUCEK, P. (2017) Digital Inheritance on Social Networks – Young People's Perception. In: Strategic Management and its Support by Information Systems (SMSIS). Ostrava,: VŠB – Technical University of Ostrava pp. 453–460. ISBN 978-80-248-4046-8. ISSN 2570-5776. WOS:000417344100051

SUDZINA, F., PAVLÍČEK. A. (2017) Do Gender and Personality Traits Influence Visits of and Purchases at Deal Sites. In: IDIMT-2017 Digitalization in Management, Society and Economy, Linz: Trauner Verlag Universität, pp. 137–146. ISBN 978-3-99062-119-6. WOS:000419365600021

BENCHMARKING OF ANOMALY DETECTION ALGORITHMS ON AUTOMATED PASSWORD ATTACKS

Pavel Strnad, Lukáš Švarc

Department of Information and Knowledge Engineering
University of Economics, Prague
pavel.strnad@vse.cz, lukas.svarc@vse.cz

Keywords

anomaly detection, machine learning, password attacks, BigML

Abstract

This paper focuses on benchmarking several supervised and unsupervised anomaly detection algorithms. Experiments are performed on KDDCUP'99 dataset using BigML platform, specifically on automated password attacks on Telnet protocol. Results of each algorithm in this paper are discussed and compared to each other as well as other researchers results. In conclusion, supervised techniques are considered superior to unsupervised in case of automated password attacks.

1. Introduction

Due to increasing computing power of processors and network capacity, potential of automated network attacks is growing enormously in recent years. Nowadays even smart light bulb, fridge, television or other IoT device with poor security is vulnerable to being hacked and misused as a resource to perform automated network attacks over the internet (Abomhara, 2015).

There are several types of attacks, ranging from simple denial of service to advanced identity thefts using social engineering (Hoque, 2014). Not all of them are possible to be automated. Most popular automated attacks can be divided into three groups. First of them is DDoS – Distributed Denial of Service attacks, where network devices under attacker's control focus one particular server or system and try to make it unavailable. Second group of automated attacks is searching for known vulnerabilities of devices reachable via internet and tries to exploit them. That's why it's so important to keep all system updated. Third group focuses on password attacks using primarily brute-force methods based on either dictionary or random input methods (Chaboya et al., 2006). Many network services, systems and devices reachable via internet use password-based authentication. Protocols such as Telnet, SSH and HTTP/S are most popular methods for user authentication over network.

Obviously, not all traffic, exploit or password input must be targeted attack. To distinguish one, it's not an easy task. Anomaly detection algorithms play a huge part in it. Main objective of this paper is to distinguish automated password attacks from legitimate user logins by using specific supervised and unsupervised anomaly detection methods.

2. Anomaly Detection Algorithms

Last year we prepared a large overview of recent anomaly detection approaches in computer networks and compared several supervised and unsupervised algorithms on theoretical basis (Svarc et al.,2019). Following that research, we chose four anomaly detection techniques which suited best this task, benchmarked, compared and discussed them.

K-means is a classical and well-known unsupervised clustering algorithm. In this algorithm, after an initial random assignment of data points to K clusters, the centres of clusters are computed and the data points are allocated to the clusters with the closest centres. The process is repeated until the cluster centres do not significantly change. Once the cluster assignment is fixed, the mean distance of a data point to cluster centres is used as the score. K-means clustering algorithm is well known data mining algorithms that can be used in anomaly detection. It has been used in an attempt to detect anomalous user behavior, as well as unusual behavior in network traffic. As the algorithm iterates through the dataset, each cluster's architecture is updated. In updating clusters, data points are removed from one cluster and added to another. The updating of clusters causes the values of the centroids to change. This change is a reflection of the current cluster datapoints. Once there are no changes to any cluster, the training of the k-means algorithm is complete. At the end of the k-means algorithm, the „k" cluster centroids are created, and the algorithm is ready for classifying traffic (Zhengxim, 2001).

Isolation Forest or iForest is an unsupervised algorithm that builds an ensemble of iTrees for a given dataset, then anomalies are those instances which have short average path lengths on the iTrees. There are only two variables in this method: the number of trees to build and the sub-sampling size. iForest's detection performance converges quickly with a very small number of trees, and it only requires a small sub-sampling size to achieve high detection performance with high efficiency. As a result, iForest has a linear time complexity with a low constant and a low memory requirement which is ideal for high volume data sets. Essentially, Isolation Forest is an accurate and efficient anomaly detector especially for large databases. Its capacity in handling high volume databases is highly desirable for real life applications (Liu et al., 2008).

Deepnets belong to supervised algorithms and are part of a broader family of classification and regression methods based on learning data representations from a wide variety of data types (e.g., numeric, categorical, text, image). Deepnets have been successfully used to solve many types of classification and regression problems in addition to social network filtering, machine translation, bioinformatics and similar problems in data-rich domains (Cetinsoy, 2017). Deep learning is an area of machine learning which applies neuron-like structure for learning tasks. Deep learning has profoundly changed the way we approach learning tasks by delivering monumental progress in different disciplines like speech processing, computer vision, and natural language processing to name a few. It is only relevant that this new technology must be investigated for information security applications. Naseer and his team under IEEE investigated the suitability of deep learning approaches for anomaly-based intrusion detection system. Their research showed exceptional performance with 85% and 89% accuracy on test dataset which demonstrates the fact that Deep learning is not only viable but rather promising technology for information security applications like other application domains (Naseer et al.,2018).

Logistic Regression is an example of supervised algorithms and works very similar to linear regression, but with a binomial response variable. The greatest advantage is the fact that you can use continuous explanatory variables and it is easier to handle more than two explanatory variables simultaneously. Although apparently trivial, this last characteristic is essential when we are interested in the impact of various explanatory variables on the response variable. The main advantage is to avoid confounding effects by analyzing the association of all variables together (Sperandei, 2014).

3. Dataset

During the last decade, anomaly detection has attracted the attention of many researchers to overcome the weakness of signature-based IDSs in detecting novel attacks. KDDCUP'99 is the mostly widely used dataset for the evaluation of these systems (Tavallaee et al.,2009). The 1998 DARPA Intrusion Detection Evaluation Program was prepared and managed by MIT Lincoln Labs. The objective was to survey and evaluate research in intrusion detection. A standard set of data to be audited, which includes a wide variety of intrusions simulated in a military network environment, was provided. The 1999 KDD intrusion detection contest uses a version of this dataset. Lincoln Labs set up an environment to acquire nine weeks of raw TCP dump data for a local-area network (LAN) simulating a typical U.S. Air Force LAN. They operated the LAN as if it were a true Air Force environment but peppered it with multiple attacks. The raw training data was about four gigabytes of compressed binary TCP dump data from seven weeks of network traffic. This was processed into about five million connection records. Similarly, the two weeks of test data yielded around two million connection records. A connection is a sequence of TCP packets starting and ending at some well-defined times, between which data flows to and from a source IP address to a target IP address under some well-defined protocol. Each connection is labeled as either normal, or as an attack, with exactly one specific attack type. Each connection record consists of about 100 bytes. Attacks fall into four main categories:

DOS – denial of service is an attack category, which depletes the victim's resources thereby making it unable to handle legitimate requests – e.g. syn flooding. Relevant features: "source bytes" and "percentage of packets with errors"

R2L – unauthorized access from a remote machine, the attacker intrudes into a remote machine and gains local access of the victim machine. E.g. password guessing. Relevant features: network level features – "duration of connection" and "service requested" and host level features - "number of failed login attempts"

U2R – unauthorized access to local super user (root) privileges is an attack type, by which an attacker uses a normal account to login into a victim system and tries to gain root/administrator privileges by exploiting some vulnerability in the victim e.g. buffer overflow attacks. Relevant features: "number of file creations" and "number of shell prompts invoked"

Probing – surveillance and other probing attack's objective is to gain information about the remote victim e.g. port scanning. Relevant features: "duration of connection" and "source bytes"

Each dataflow contains 41 discrete or continuous values such as duration, service type, transferred bytes etc. Each dataflow is also labeled as normal traffic or as a one of twenty-two specific types of attack, each of them belonging to one of the four categories mentioned above (Stolfo et al., 2000).

4. Experiments

In our experiments we focused on automated password attacks on Telnet protocol. Telnet is a network protocol that provides two-way interactive communication compatibility for computers on the internet and local area networks. Telnet has a command-line interface and is famous for being the original protocol from when the internet first launched. In time, Telnet's use declined in favor of SSH (Secure Shell or Secure Socket Shell) due to serious security concerns when it was used over an open network. Telnet lacks authentication policies and data encryption. Nevertheless, password attacks on Telnet are very similar, if not exactly same, to attacks on SSH or HTTP/S. We can consider Telnet testing adequate. In KDDCUP'99 dataset Telnet traffic consists of 94.54% traffic identified as normal, while 5.46% is identified as an attack. Out of forty-four available fields, we observed four

of them, which are important in detecting automated password attacks: duration (Field1), dst_bytes (Field6), num_failed_logins (Field11), logged_in (Field12).

Table 1 – Used Fields of KDDCUP'99 Dataset

Field	Type	Description
duration (1)	continuous	length (number of seconds) of the connection
dst_bytes (6)	continuous	number of data bytes from destination to source
num_failed_logins (11)	continuous	number of failed login attempts
logged_in (12)	discrete	1 if successfully logged in; 0 otherwise

We run experiments on BigML framework, which is robust cloud-based platform that offers Machine Learning as a Service (MLaaS). It supports dozens of modifiable supervised and unsupervised learning techniques. It offers multi-user access, academical licensing, modern graphical user interface and interactive visualization of results. We performed series of experiments using selected algorithms and modified their variables.

K-means was first algorithm to run. It is very popular unsupervised clustering algorithm, which is quite easy to implement. After initial testing, K value was set to 21 and it was run over prepared KDDCUP'99 dataset. Clusters containing all data were created as you can see in Figure 1. Black circle cluster contains 78,79% instances of attacks. Although, the detection rate might seem quite high, there is a problem with identifying correct cluster which represent attack scenario. Even after fine tuning, clusters are quite close to each other and it's fairly difficult to recognize the one containing anomalies.
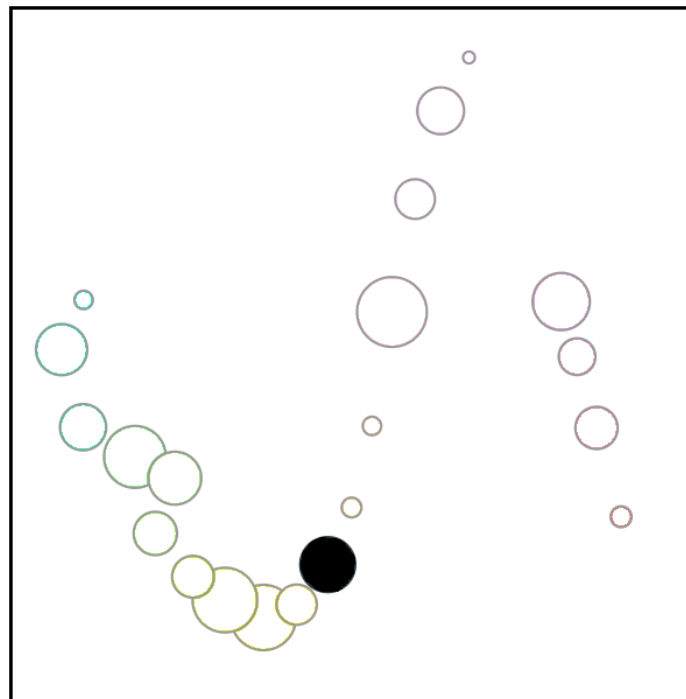
Figure 1 - K-means Generated Clusters

Isolation Forest was second unsupervised algorithm to experiment with. After initial testing, Forest size was set to 97. It reached detection rate of 38,27%. Its score is observable in Figure 2. Because KDDCUP'99 dataset contains many Telnet attacks, which are identical in most parameters, large circle of detected attacks is noticeable.
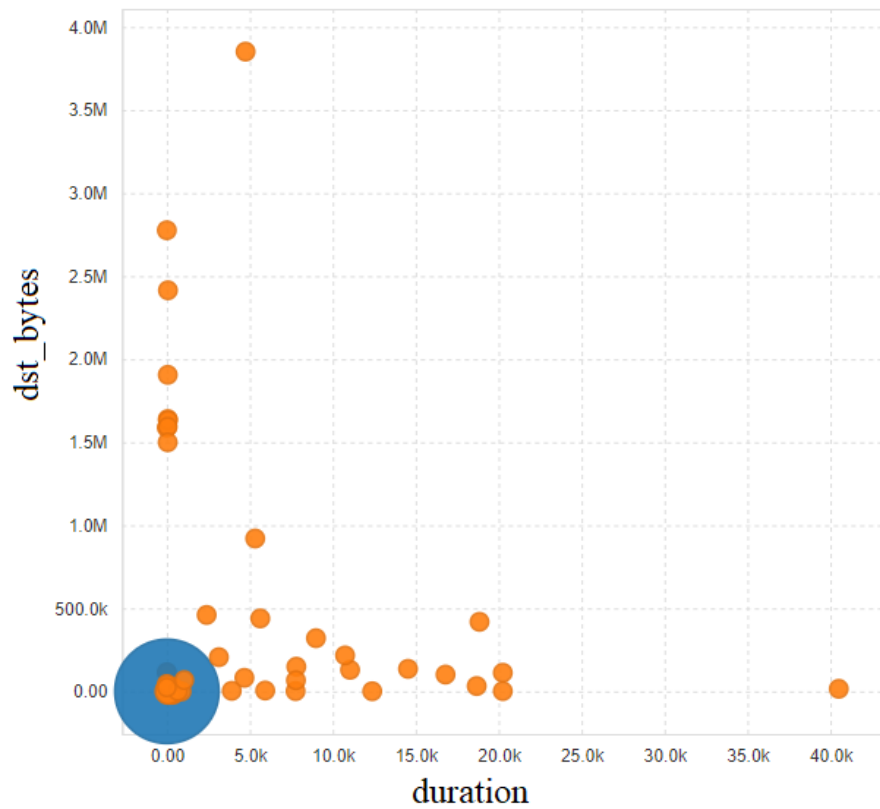
Figure 2- Isolation Forest Results

Deepnets was first supervised algorithm to run. Train and Test data were prepared by Kaggle platform based on KDDCUP'99 dataset (Anush, 2019). Based on training data it generated field importance analysis and weighted duration to 61.61%, dst_bytes to 13.39%, num_failed_logins to 19%, logged_in to 6%. In testing phase, it predicted 97,44% attacks and reached detection rate of 88.23%.

Table 2 – Deepnets Ratio of True/False Positive/Negative

Actual Predicted	Actual Password attack	Actual Normal traffic	Actual
Predicted Password attack	5.32%	0.14%	5.46%
Predicted Normal traffic	0.71%	93.83%	94.54%
Predicted	6.03%	93.97%	100%

Logistic Regression was second supervised, and last algorithm overall, to run. In testing phase, it predicted 94.87% attacks and reached detection rate of 82.09%.

Table 3 – Logistic Regression Ratio of True/False Positive/Negative

Actual Predicted	Actual Password attack	Actual Normal traffic	Actual
Predicted Password attack	5.18%	0.28%	5.46%
Predicted Normal traffic	1.13%	93.41%	94.54%
Predicted	6.31%	93.69%	100%

5. Conclusion and future work

We performed series of experiments with selected algorithms, K-means, Isolation Forest, Deepnets and Logistic Regression using BigML framework on KDDCUP'99 dataset. Results of each method can be seen in Table 4. Unlike other researches, who usually focused on whole dataset and all intrusion events, we focused strictly on automated password attacks. In conclusion, we found out, that unsupervised learning methods, such as K-means and Isolation Forest don't provide good enough results to be considered as a potential anomaly detection algorithm to use. As we presumed, supervised learning methods provided much better results. Deepnets algorithm reached detection rate 88.23%, while Logistic Regression reached detection rate of 82.09%.

Table 4 – Results of Anomaly Detection Algorithms on KDDCUP'99 Dataset Telnet Traffic

Method	Category	Detection rate
Deepnet	Supervised	88.23%
Logistic Regression	Supervised	82.09%
K-means	Unsupervised	78.79%
Isolation Forest	Unsupervised	38.27%

Although KDDCUP'99 dataset is quite old, it provides good enough results for evaluation and proof of concept purposes. There have been several papers published using KDDCUP'99 dataset testing anomaly detection methods, but most of them focused on all intrusion events, not specific cases such as password attacks. If we compare our results to other researchers, Joshi and his coworkers tested Hidden Markov Model (HMM) on same dataset and reached 79% detection rate (Joshi et al., 2005). Kim and his coworkers tested Fusions of GA and SVM for Anomaly Detection on same dataset and reached 98% detection rate (Seong et al., 2005). Chandrashekhar and his coworkers experimented with clustering techniques on the same dataset and with K-means algorithm best case achieved detection rate of 91.02%. (Chandrashekhar et al., 2012). We can conclude that password attacks are more difficult to detect, which explains lower detection rate when testing password attacks only compared to all intrusion events.

In conclusion, we consider supervised learning methods superior to unsupervised learning methods in case of automated password attack in computer networks. As such, we came to the same conclusion as other researchers who worked with all intrusion events contained in KDDCUP'99 dataset. For example, Laskov and his coworkers say that their experiments demonstrate that the supervised learning methods significantly outperform the unsupervised ones if the test data contains no unknown attacks (Laskov et al., 2005).

In the future, we would like to focus primarily on supervised learning methods, possibly implement new modified version of well-known algorithms. We would like to use newer and more complex datasets, such as CSE-CIC-IDS2018. However, we need to overcome limitation of BigML platform which allows only 4 GB size of datasets with academic license, because CSE-CIC-IDS2018 dataset has the size of about 200 GB.

6. Acknowledgment

7. References

Abomhara, M., & and Kø Ien, G. M. (2015). Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks. In: Journal of Cyber Security and Mobility 4, no. 1: 65–88. https://doi.org/10.13052/jcsm2245-1439.414

Anush (2019). Kaggle: Network anomaly detection. https://www.kaggle.com/anushonkar/network-anamoly-detection

Cetinsoy, A. (2017). Introduction to Deepnets. In: The Official Blog of BigML.Com. https://blog.bigml.com/2017/09/26/introduction-to-deepnets/

Chaboya, D. J., Raines, R. A., Baldwin R.O. & Mullins B. E. (2006). Network Intrusion Detection: Automated and Manual Methods Prone to Attack and Evasion. In: IEEE Security Privacy 4, no. 6, 36–43. https://doi.org/10.1109/MSP.2006.159.

Chandrashekhar, A.M. & Raghuveer K. (2012). Performance evaluation of data clustering techniques using KDD Cup-99 Intrusion detection data set. In: International Journal of Information & Network Security (IJINS) Vol.1, No.4, pp. 294~305, ISSN: 2089-3299

Hoque, N., Bhuyan M. N., Baishya, R. C., Bhattacharyya, D. K., & Kalita, J. K. (2014) Network Attacks: Taxonomy, Tools and Systems. In: Journal of Network and Computer Applications 40: 307–24. https://doi.org/10.1016/j.jnca.2013.08.001

Joshi, S. S., & Phoha V. V. (2005). Investigating Hidden Markov Models Capabilities in Anomaly Detection. In: Proceedings of the 43rd Annual Southeast Regional Conference - Volume 1, 98–103. ACM-SE 43. Kennesaw, Georgia: Association for Computing Machinery. https://doi.org/10.1145/1167350.1167387

Laskov, P., Düssel, P., Schäfer C., & Rieck, K. (2005). Learning Intrusion Detection: Supervised or Unsupervised. In: Image Analysis and Processing – ICIAP 2005, 50–57. Berlin, Heidelberg: Springer. https://doi.org/10.1007/11553595_6

Liu, F. T., Ting, K. M., & Zhou, Z. (2008). Isolation Forest, Eighth IEEE International Conference on Data Mining, Pisa, 2008, pp. 413-422.

Naseer, S., Saleem Y., Khalid, S., Bashir, M. K., Han J., Iqbal, M. M., & Han, K. (2018). Enhanced Network Anomaly Detection Based on Deep Neural Networks. In: IEEE Access 6, 2018: 48231–46. https://doi.org/10.1109/ACCESS.2018.2863036

Seong, K. D., Nguyen, H., Ohn S., & Park J. S. (2005). Fusions of GA and SVM for Anomaly Detection in Intrusion Detection System. In: Advances in Neural Networks – ISNN 2005, 415–20. Berlin, Heidelberg: Springer. https://doi.org/10.1007/11427469_67

Sperandei, S. (2014). Understanding Logistic Regression Analysis. In: Biochemia Medica 24, no. 1, 12–18. https://doi.org/10.11613/BM.2014.003.

Stolfo, S., Fan, W., Lee, W., Prodromidis A., & Chan P. (2000). Cost-Based Modeling and Evaluation for Data Mining with Application to Fraud and Intrusion Detection: Results from the JAM Project. In: Proceedings DARPA Information Survivability Conference and Exposition. DISCEX'00, 2:130–44 vol.2. https://doi.org/10.1109/DISCEX.2000.821515

Svarc, L., & Strnad, P. (2019). Recent Anomaly Detection Approaches in Computer Networks. In: IDIMT 2019: Innovation and Transformation in a Digital World - 27th Interdisciplinary Information Management Talks, 229–36.

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A Detailed Analysis of the KDD CUP 99 Data Set. In: 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 1–6, 2009. https://doi.org/10.1109/CISDA.2009.5356528

Zhengxim, C. (2001). Data Mining and Uncertain Reasoning: An integrated approach. Willey, New York, 392 pages, ISBN 0471388785

ESTABLISHING AN AVATAR-BASED PRESENCE ON INSTAGRAM

Helena Lovasz-Bukvova, Roger Hage, Martin Waiguny, Tina Gruber-Mücke

IMC FH Krems, Krems an der Donau, Austria

helena.lovasz-bukvova@fh-krems.ac.at, roger.hage@fh-krems.ac.at,
martin.waiguny@fh-krems.ac.at, tina.gruber-muecke@fh-krems.ac.at

Keywords

Social Media, Instagram, avatar, Social-Media marketing, self-presentation

Abstract

Instagram provides a space for a digitalized self-presentation of individuals as well as organizations. Being image-based, it creates an illusion of credibility and intimacy thorough a connection to reality, such as places, people, or events. However, not all Instagram users might have a real individual or organization in the background – public campaigns, for example, might choose to act independently of an explicit backer. In this paper, we explore the possibility of creating a purely artificial, avatar-based presence on Instagram. Our findings indicate that while an avatar can gain visibility and presence, it may require greater investment to engage followers in a reciprocal interaction.

1. Introduction

Among the available social-media providers, Instagram, a photo-sharing platform, has been one of the fastest-growing platforms within the last 5 years, becoming a relevant marketing tool for companies and organizations (Měsíček & Petrus, 2019) but also for information sharing. With its visual stories, Instagram differs from traditionally text-oriented platforms (e.g. Facebook, Twitter) in its reliance on image sharing. The sharing of photos – typically representing the captures of offline events, objects, or places – Instagram uses the digitalization of real experiences to create an online presence as an extension of an offline presence. Even though the digitalization process need not be always authentic, there is an illusion of reality: the idea of an actual person or an actual organization acting and living behind a profile. Indeed, both compelling content and an attractive creator are essential to creating a parasocial relationship between an influencer and the crowd (Lockie, 2019).

However, in some cases, there may not be a recognizable creator behind an Instagram account, such as in initiatives that focus on a particular message but are not backed up by any established organization. These may find it difficult to give their virtual presence an identity. This paper explores the challenges of creating an Instagram presence independent of a personal or organizational representative, based solely on an avatar with an abstract identity. For this purpose, an Instagram account for an e-mobility campaign was created and operated for 6 months to find out *whether it is possible to establish an avatar-based internet presence as an Instagram influencer.*

In the following, we first discuss existing research on organizational self-presentation on WWW and especially on Instagram (section 2). In section 3, we describe the design of the case study and analyze

the account metrics during the deployment of the campaign. In the conclusions (section 4), we summarize the findings and lessons learned.

2. Literature review

Communication and social interaction are an important part of human behaviour. The WWW, especially the later development of Web 2.0 platforms such as SNS (social networking services), is meant to support such behaviour in the digital world. SNS provide users with a variety of interaction opportunities and roles (Kumar, Novak, & Tomkins, 2010), facilitating a rich exchange that complements offline interaction and relationship building (Hampton & Wellman, 2000; Kim et al., 2019).

2.1. Impression management on Instagram

Digital interaction, although using new kinds of communication channels, has been shown to follow patterns known from offline interaction. Personal digital networks tend to mirror offline contacts and relationships, providing opportunities for new interaction, in addition to offline communication (Hampton & Wellman, 2000; Kim et al., 2019). Similar to interaction in the offline world, and even more so in the virtual world, participants engage in an active creation of "impression management" – adjusting their behavior and communication to create a particular image (Goffman, 1959; Rosenfeld, Giacalone, & Riordan, 1995).

Harrison and Budworth (2015) note that online impression management – again mirroring offline behavior – has a verbal and a non-verbal dimension. The verbal dimension is represented through explicit messages (e.g. texts), while the non-verbal relies on visual clues on photographs or the design of private web spaces. This points to the relevance of imagery in an online presentation. Active impression management is even more important in case of organizational virtual presence; because organizations have no real personality, their presence on webpages or platforms can be considered fully constructed (Winter, Saunders, & Hart, 2003).

Instagram users name many reasons why they use Instagram, such as (Sheldon & Bryant, 2016) surveillance and knowledge about others, documentation, coolness, and creativity or social interaction, archiving, self-expression, escapism, and peeking (Lee, Lee, Moon, & Sung, 2015). Instagram appears to be particularly suited for online impression management. Relying mainly on sharing images, Instagram creates an illusion of reality and hence can be perceived as more credible (Sundar, 2008) and more intimate (Pittman & Reich, 2016) than text-based platforms. Organizations use Instagram to present their more "human" side, connected to humor and emotions (Ginsberg, 2015).

2.2. Credibility and trust on Instagram

The illusion of reality created by the visual media on Instagram would be expected to lead to greater credibility and trustworthiness of Instagram personas. The following is a summary of research-based recommendations for a credible and trustworthy Instagram content presentation.

Believability: The conveyed communication must not be perceived as insincere ("too much") (Pollach, 2005) and be in aligned to other communications even outside of Instagram (Van-Tien Dao, Nhat Hanh Le, Ming-Sung Cheng, & Chao Chen, 2014).

Compelling message: The message must be of interest and value to the followers, providing an informative or entertainment value (Lou & Yuan, 2019; Pollach, 2005; Van-Tien Dao et al., 2014).

Advocacy: Peer influence plays an important part on Instagram, making "liked" or indorsed images more likely to be considered interesting ("bandwagon cues") (McGlynn, Zhou, Han, & Huang, 2019; Sherman, Greenfield, Hernandez, & Dapretto, 2018).

Public activity: Regular and frequent posting activity makes the account seem "alive" and hence real and credible (McGlynn et al., 2019).

Positive interaction: Positive interaction, such as liking, praise, and positive feedback leads to the strengthening of relationships (Orben & Dunbar, 2017). However, Orben and Dunbar (2017) also note that while Instagram presence and interaction can create an illusion of intimacy and relationship, they lack certain aspects of a real relationship. One of the missing features is the slow growth of intimacy by the participants in the relationship. Therefore, even positive feedback can be perceived as inappropriate, if it considered as being (suddenly) too intimate and hence lead to a trust loss.

The persona of the influencer: Research into parasocial relationships, further highlights the importance of the person of influencer (e.g. Farivar, Wang, & Yuan, 2019). Parasocial relationships are one-sided (e.g. influencer and followers), but the larger crowd feels a form of connectedness which is not perceived by the other side (Perse & Rubin, 1989). Lockie (2019) found that in particular personal information contributes to parasocial connectedness for blogs and vlogs. In the case of Instagram, the visual attractiveness of the influencer positively relates to a parasocial relationship (Sokolova & Kefi, 2020). Furthermore, being authentic as an influencer (Audrezet, de Kerviler, & Guidry Moulard, 2018) and congruence between the influencer and the posted information (Shan, Chen, & Lin, 2019) also strengthens parasocial relationships.

2.3. Avatar-based Instagram persona

Avatars are used to personify objects, emotions or processes and are a special form of anthropomorphism. Prior research found avatars to be successful in gaining attention and nurture perception (Nowak & Rauh, 2005). In particular comic-like avatars fast-track the creation of a positive or negative empathic relation (Gardner, Herman, & Keen, 2011). While real people can use an avatar as a digital representation of themselves, avatars can also have their own personality. When considering the use of a person-independent avatar on Instagram, the previous research summary suggests that such a profile may be limited in its establishing a credible and trustworthy presence on Instagram. Although it can still follow the recommendations regarding the published content (believable content, compelling message, public activity, positive interaction), it may face difficulties due to a lacking personability (persona, advocacy).

3. Method

To judge the feasibility of establishing an avatar-based presence as an Instagram influencer, we posed two research questions:

RQ1: Can an avatar-based Instagram presence establish a presence and gain visibility on Instagram?

RQ2: Can an avatar-based Instagram presence create a reciprocal relationship with its followers?

The research method was a case study in which an actual Instagram avatar-account was created and operated for 6 months. The avatars presence and gain were judged by the number of gained followers as well as the reach of advertising campaigns. The existence of a reciprocal relationship was operationalized as the intensity of likes and comments on the avatar's posts.

3.1. Case study: Poldi

In a project concerned with promoting e-mobility in Austria, a campaign was launched to provide citizens with a better understanding of e-cars and e-mobility. For this purpose, an Instagram channel was created that was intended to providing adequate content in form of visuals and authentic emotional stories of everyday life to refute the myths surrounding e-mobility and post about currently relevant topics in society. While supported by the Climate-Energy Fund and run by the IMC FH Krems, the campaign was meant to be independent and not be considered as an explicit outreach of either of these institutions. To prevent a feeling of manipulation, an avatar personality of an e-car called Poldi (see Figure 1) was created with a background story and "personal" information provided on a dedicated landing page (https://www.epoldi.at/). And as highlighted before, congruence between topic and person is a key element in impression management as anthropomorphism of an e-car itself was chosen. As homophily and androgyny are essential factors (Novak & Rauh 2005) in avatars attractiveness it was decided to draw the avatar with some masculine traits, whereas in all textual information a female address was used. Furthermore, to relate to visible e-cars on the road, the fictitious character incorporates similar features from major European, Asian and American car brands to allow to relate quickly to the character. The elaborate creation of the avatar was meant to personify the channel, connecting it to – albeit not real – but a person-like persona.
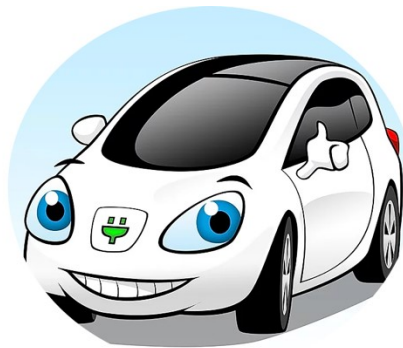
Figure 1: Avatar Poldi

The landing page and the Instagram account were created on July 1, 2019. The avatar Instagram interaction had the form of an information campaign that ran until December 31, 2019, posting 2-3 times a week. The posts represented a mixture of informative 1:n communication, reflective posts for seeding interaction, and emotional/fun posts meant to personify the avatar: (1) *informational posts* (posted on Wednesdays, they explain and refute myths about e-cars, following a standardized design always including a large picture and a question), (2) *reflective posts* (Sustainable Saturday posts, inviting people to reflect on the sustainability effects of clean energy, asking them to debate the topic), (3) *emotional posts* (posted irregularly, they address a current topic, e.g. women's day, talk-like-a-pirate day, or current events, e.g. parliamentary elections, tradeshows related to e-mobility, or just small anecdotes and memes), and (4) *seasonal posts* (before Christmas, 24 daily posts were published as an "advent calendar").

Furthermore, additional commercial advertising was used throughout the campaign to support follower acquisition. Commercial advertising was directed towards the profile to gain followers, but also presented posts mentioned above to make the impact comparable. The promotions target in particular people interested in cars and sustainability, but also some other interests like hobbies addressed in the posts, or current hot topics, to gain new followers respectively expose people, who did not consume e-mobility information before to the content.

3.2. Analysis

From July 1, 2019, until December 31, 2019, the Instagram presence has garnered 1048 followers. Figure 2 shows that the number of followers had risen fairly constantly, with patterns of steep growth that coincide with the advertising campaigns. The small loss of followers after 90 days was caused by Instagram removing fake advertising accounts. The growth ceased once the advertising campaigns stopped, but the number of followers remained steady from thereon. After the 6-month-campaign, the account continued posting content, but with a much lower frequency than before, yet this did not result in the loss of followers.
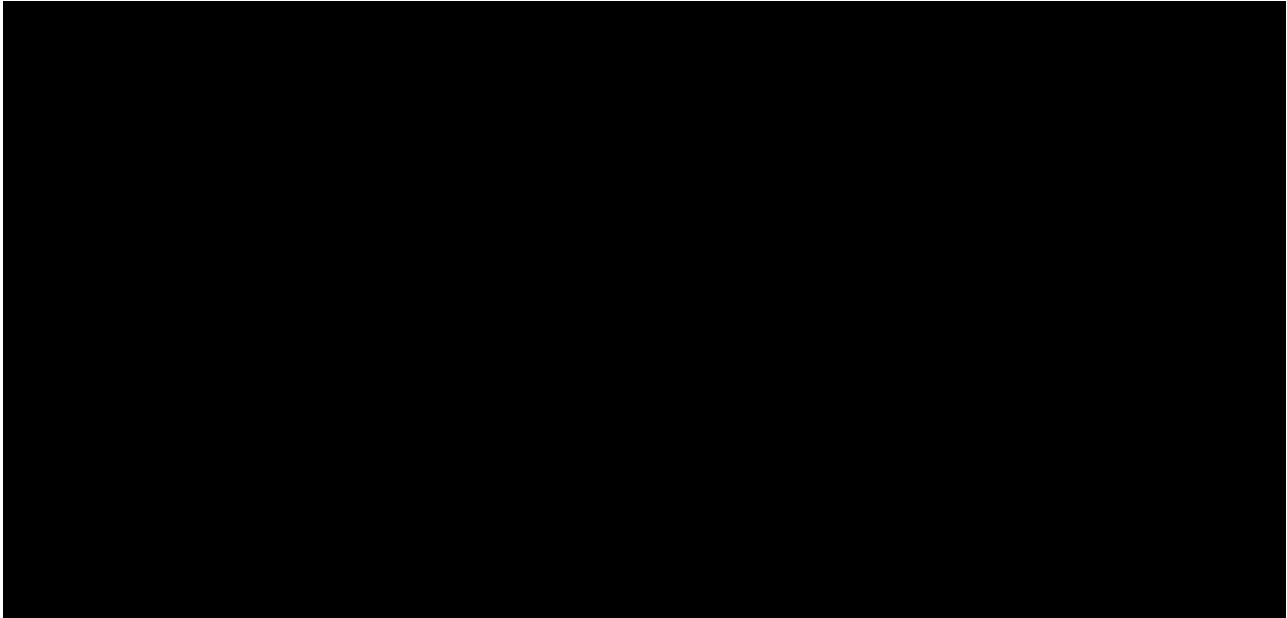
Figure 2: Growth of followers

Looking into the commercial advertising performance, promotions between 29th of July and 26th of December were launched, which generated 59261 gross impressions, with an average spent of 4,80 Euro per 1000 impressions. Given the promotions only ran in Austria, using the German language, we achieved 414 direct reactions (divert to the profile) to the posts by forwarding to the profile. We promoted 6 different categories of posts (see Table 1). In line with the in section 2.2 mentioned recommendations, humor and real-life promoted posts performed best showing that entertainment and authenticity are also important for a completely fictitious avatar. There was also a steady reach and impressions for each post which is between 250 and 500. This appears to agree with the expected effects suggested by the recommendations for creating a credible and trustworthy presence (see section 2.2).

Table 1: promotions performance

Category	Impressions	Spent	Reaction	Impressions /Spent	Impressions /Reaction	Reactions /Spent
Myths (about e-cars)	13891	66	119	210,47	116,73	1,80
Current (relates to something important in society in the period)	12323	74	113	166,53	109,05	1,53
How to e-car (practical tips)	11686	74	72	157,92	162,31	0,97
Humor (fun videos and posts)	11331	34	37	333,26	306,24	1,09
Facts (about eMobility)	4454	22	25	202,45	178,16	1,14
Real Life (shows episodes of real e-car use)	5039	15	50	335,93	100,78	3,33

Although the account managed to gain visibility and assume a (moderate) presence, the level of interaction and hence influence of the channel appears to be limited (see Table 2). While there were on average 19.84 likes per post, which does show a level of engagement with the audience, the average number of comments was only 0.5, with the majority of posts being without any comments whatsoever.

Table 2: Influence measures (based on Segev, Avigdor, & Avigdor, 2018)

total posts	Number of posts during the campaign	92
total likes	Number of likes during the campaign	1825
total comments	Number of comments during the campaign	46
total followers	Number of followers after the campaign	1048
likes	Average likes per post	19.84
comments	Average comments per post	0.5
$\sqrt{\text{likes·followers}}$	Geometric mean of likes and comments	3.15
followers / posts	Followers per post	11.39
comments / likes	Relation of comments to likes, as likes are more likely to be influenced by advertising than comments	0.03

In summary, during the campaign, the posting and commenting activities were regular and frequent. The steady rise and the loyalty of followers suggest that the message was sufficiently compelling and of interest, but without eliciting reciprocal reaction such as likes and comments. Likes are easier to boost through advertising, hence the comparatively high number of average likes per post cannot be considered a sole measure of interaction. It is the lack of comments which points to the lack of reciprocity in the user base, even though some posts were specifically geared towards creating engagement and discussion. This could be in part due to the channel language (German-speaking audience is much smaller than English-speaking one), though this was not a hindrance in gaining base visibility. Hence, it can point to the unwillingness of the audience to engage with an obviously unreal avatar personality.

4. Conclusion

The purpose of this paper was to analyse the visibility and relationships of an avatar-based Instagram presence based on a case study of a public information campaign on e-mobility in Austria. The avatar Poldi, an imaginary comic-like e-car, was able to create a visible Instagram presence as a result of a six-month campaign, gaining and maintaining over 1000 followers. This shows that it is possible to create an avatar-based influencer with substantial followership. Engaging with the followers a meaningful reciprocal interaction proved to be more challenging for the avatar. There can be several reasons for this: the topic (e-mobility is not very emotionally charged), content (focus on informing rather than conversation), or target group (German-speaking audience). Another reason could be that the missing reality behind the avatar makes it more difficult to create the parasocial relationship that is seen to exist for human Instagram influencers. Possibly, the use of solely avatar-based Instagram profiles requires a greater investment in the virtual impression management and might even need explicit advocacy of real human or organisational influencers.

5. Acknowledgements

6. References

Audrezet, A., de Kerviler, G., & Guidry Moulard, J. (2018). Authenticity under threat: When social media influencers need to go beyond self-presentation. Journal of Business Research. https://doi.org/10.1016/j.jbusres.2018.07.008

Farivar, S., Wang, F., & Yuan, Y. (2019). Meformer vs. informer: Influencer type and follower behavioral intentions. 25th Americas Conference on Information Systems, AMCIS 2019.

Gardner, J., Herman, D., & Keen, S. (2011). Fast Tracks to Narrative Empathy: Anthropomorphism and Dehumanization in Graphic Narratives. SubStance, 40(1), 135–155. https://doi.org/10.1353/sub.2011.0003

Ginsberg, K. (2015). „Instabranding : Shaping the Personalities of the Top Food Brands on Instagram". Journal of Undergraduate Research in Communications, Vol. 6(1), 78–91.

Goffman, E. (1959). The Presentation of Self in Everyday life. New York, NY, USA: Anchor Books.

Hampton, K. N., & Wellman, B. (2000). Examining community in the digital neighborhood: Early results from Canada's wired suburb. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1765 LNCS, 194–208. https://doi.org/10.1007/3-540-46422-0_16

Harrison, J. A., & Budworth, M. H. (2015). Unintended consequences of a digital presence: Employment-related implications for job seekers. Career Development International, 20(4), 294–314. https://doi.org/10.1108/CDI-06-2014-0080

Kim, Y. C., Shin, E., Cho, A., Jung, E., Shon, K., & Shim, H. (2019). SNS Dependency and Community Engagement in Urban Neighborhoods: The Moderating Role of Integrated Connectedness to a Community Storytelling Network. Communication Research, 46(1), 7–32. https://doi.org/10.1177/0093650215588786

Kumar, R., Novak, J., & Tomkins, A. (2010). Structure and Evolution of Online Social Networks. In P. S. Yu, J. Han, & C. Faloutsos (Hrsg.), Link mining: Models, algorithms, and applications (S. 337–357).

Lee, E., Lee, J. A., Moon, J. H., & Sung, Y. (2015). Pictures Speak Louder than Words: Motivations for Using Instagram. Cyberpsychology, Behavior, and Social Networking, 18(9), 552–556. https://doi.org/10.1089/cyber.2015.0157

Lockie, M. A. (2019). In Vlogs We Trust Consumer Trust in Blog and Vlog Content. Auckland University of Technology.

Lou, C., & Yuan, S. (2019). Influencer Marketing: How Message Value and Credibility Affect Consumer Trust of Branded Content on Social Media. Journal of Interactive Advertising, 19(1), 58–73. https://doi.org/10.1080/15252019.2018.1533501

McGlynn, J., Zhou, Y., Han, A., & Huang, Y. (2019). Present, endorsed, and active: Instagram cues that predict trust. Communication Research Reports, 36(4), 338–348. https://doi.org/10.1080/08824096.2019.1670154

Měsíček, L., & Petrus, P. (2019). Qualitative study of social media use by Small enterprises as illustrated on the Czech-German border. Proceedings of 27th Interdisciplinary Information Management Talks, IDIMT 2019, 259–266. Kutna Hora, Czechia.

Nowak, K. L., & Rauh, C. (2005). The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. Journal of Computer-Mediated Communication, 11(1). https://doi.org/10.1111/j.1083-6101.2006.tb00308.x

Orben, A. C., & Dunbar, R. I. M. (2017). Social media and relationship development: The effect of valence and intimacy of posts. Computers in Human Behavior, 73, 489–498. https://doi.org/10.1016/j.chb.2017.04.006

Perse, E. M., & Rubin, R. B. (1989). Attribution in Social and Parasocial Relationships. Communication Research, 16(1), 59–77. https://doi.org/10.1177/009365089016001003

Pittman, M., & Reich, B. (2016). Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. Computers in Human Behavior, 62, 155–167. https://doi.org/10.1016/j.chb.2016.03.084

Pollach, I. (2005). Corporate self-presentation on the WWW: Strategies for enhancing usability, credibility and utility. Corporate Communications: An International Journal, 10(4), 285–301. https://doi.org/10.1108/13563280510630098

Rosenfeld, P., Giacalone, R. A., & Riordan, C. A. (1995). Impression Management in Organizations: Theory, Measurement, Practice. London, UK: Routledge.

Segev, N., Avigdor, N., & Avigdor, E. (2018). Measuring Influence on Instagram: a Network-oblivious Approach. http://arxiv.org/abs/1806.00881

Shan, Y., Chen, K.-J., & Lin, J.-S. (Elaine). (2019). When social media influencers endorse brands: the effects of self-influencer congruence, parasocial identification, and perceived endorser motive. International Journal of Advertising, 1–21. https://doi.org/10.1080/02650487.2019.1678322

Sheldon, P., & Bryant, K. (2016). Instagram: Motives for its use and relationship to narcissism and contextual age. Computers in Human Behavior, 58, 89–97. https://doi.org/10.1016/j.chb.2015.12.059

Sherman, L. E., Greenfield, P. M., Hernandez, L. M., & Dapretto, M. (2018). Peer Influence Via Instagram: Effects on Brain and Behavior in Adolescence and Young Adulthood. Child Development, 89(1), 37–47. https://doi.org/10.1111/cdev.12838

Sokolova, K., & Kefi, H. (2020). Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. Journal of Retailing and Consumer Services, 53, 101742. https://doi.org/10.1016/j.jretconser.2019.01.011

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. Digital Media, Youth, and Credibility, 73–100. https://doi.org/10.1162/dmal.9780262562324.073

Van-Tien Dao, W., Nhat Hanh Le, A., Ming-Sung Cheng, J., & Chao Chen, D. (2014). Social media advertising value: the case of transitional economies in Southeast Asia. International Journal of Advertising, 33(2), 271. https://doi.org/10.2501/IJA-33-2-271-294

Winter, S. J., Saunders, C., & Hart, P. (2003). Electronic window dressing: impression management with Websites. European Journal of Information Systems, 12(4), 309–322. https://doi.org/10.1057/palgrave.ejis.3000470

BIOMETRIC SYSTEMS AND THEIR USE IN SOCIAL NETWORKS

Anton Lisnik

Catholic university in Ružomberok, Faculty of education
anton.lisnik@ku.sk

Jana Janíčková, Katarína Zimmermannová

University Matej Bel, Faculty of economics
jana.janickova@umb.sk, katarina.zimmermannova@umb.sk

Keywords

Biometrics, biochip, social networks

Abstract

Biometrics is a way of identifying a specific person, which in the system allows you to identify a specific person and thus log in to the system. Common use of biometrics is part of various systems of banks, operators or the work environment. These systems require a high degree of protection with accurate identification of the person, while requiring a high degree of accuracy and security. Less demanding security is associated with the comfort of using a car or living, where the user identifies the thing, adapts the thing to him and creates an environment that suits the user of the thing. It is, for example, the setting of the temperature in the apartment, lighting or music in the car as well as the setting of the seat or steering wheel.

The most modern trend in biometrics is the biochip - a passive device under the skin, which carries a lot of information from the identification data of the owner of the biochip to his medical history and the ability to pay by card in the store.

The article focuses on the possibilities of using bioidentification in social networks. At the same time, it examines the ethics of state-of-the-art biochip devices and biometrics in relation to their use in the social networking system.

1. Current state of knowledge

Biometrics is a field of science dealing with the process of human authentication based on its physiological or behavioural characteristics. (Drahamský, 2011)

In the field of information technology, the term biometrics refers to a system of automatic recognition of persons on the basis of their characteristic behavioural and physiological features and properties. To physiological keys, resp. characteristics of a person include our anatomical data, such as fingerprints, the texture of the iris or retina of the eye, the image of the face or the dimensions of the hands. These traits are genetically implicit, but at the same time there is a potential possibility of influencing them by various environmental factors. Behavioural, is habitual characteristics of a

person include, for example, his signature, voice, dynamics of keystrokes or walking. It is therefore a characteristic of the behaviour that a person acquires during life. We therefore call them acquired or learned. (Bača, 2009)

Biometric recognition systems constitute a set of methods typically serves two purposes, namely the identification and verification of individuals. (Chirchi, 2011)

Currently, the most modern form of biometrics is the biochip. A biochip is a set of diminished microarrays that are placed on a strong substrate that allows many experiments to be executed at the same time to obtain a high throughput in less time. This device contains millions of sensor elements or biosensors. Not like microchips, these are not electronic devices. (Young-Sik Jeong, 2014)

Passwords that are more or less legible for people are more than 60 years old and represent the easiest way to authenticate the client, but they are a very weak obstacle to being misused in some way (Doucek et al., 2020). That's why web designers come up with the idea of Multifactor Verification, which is used for work emails, whether on some work or public portals. This authentication has undergone the same change where nowadays passwords or numeric codes are used, which the client receives on SMS and is valid for only a few seconds, or otherwise secure access and verification passwords, which have either time because of local validity. Today, mobile and work devices use biometric authentication mechanisms such as fingerprint or face scans. At the time of the 3D printer, even these things got to a disadvantage, because the scanners can be deceived by a perfect imitation of the author, whether the account holder or identity.

This danger shifts people to the need for a combination of verification factors, such as a facial scan, whether the iris and a comparison with the voice or archetypal gait of the account client - multimodal biometrics. However, this is very difficult to store data and to connect to a server that performs benchmarking. The latest trend is behavioural biometrics, where the client is monitored by a device that creates a model of the client's behaviour and who is difficult to learn. Thus, the device itself remains open and available to the user. This biometrics is most often used in banking, but nowadays it is expanding into the field of insurance, communication with the state, whether its bodies - e-government, or when shopping over the Internet, even employers start to consider it (Böhmová, 2015, Černý et al. 2019).

One of the ways to avoid burdening the system with processing a lot of data is the latest developed Biochiping technologies.

Human enhancement can be achieved through various ways: through internal, wearable, or external devices/materials/substances; through hardware or software (e.g., brain–machine interface); or a smart environment (e.g., digital or virtual environments). Globally, many of such emerging technologies work by augmenting a biological function of the organism that otherwise would work "normally." (Gaspar, 2019).

This could take the form of, for example, external devices, internal implants, materials/substances intake, gene therapy or others, to augment hearing, vision or cognitive capacities, target defective/inactive genes. Specifically, such activities can be defined as any "modification aimed at improving individual human performance and brought about by science-based or technology-based interventions in the human body" (Gaspar, 2019).

2. Research procedure

The objective of the survey was to obtain relevant data for the arguments about using biometric authentication in social networks.

The survey was conducted in September and October in 2019. All respondents were residents in towns and villages near the Poprad (pp) and Banska Bystrica (bb) district.

The survey involved 150 people aged 17 to 69 years, divided into groups: students (st), employees (emp), unemployees (unemp) and retired (ret). There were 75 respondents in each group in two cities. We examined whether respondents knew the functionality of biometrics, social networks and whether they were using them.

The survey was primary, qualitative. The exploratory tool was a brief structured questionnaire which contained 10 basic questions about knowledge: purpose, functionality, content and frequency of social networks usage, preference of use over other communication tools, as well as what the expected or actual impact of biometrics, social networks on their lives. The answers were prepared as closed, with the possibility of adding your own answer.

We examined questions related to the use of an email address, where all respondents had a private email address. All respects also know and use PIN verification. Therefore, we will not tabulate these results.

The distribution of respondents is based on the results of the composition of the company according to the latest population and Slovakia and the development of unemployment for the month of October 2019. Distribution: 13 students, 25 employed, 12 unemployed and 25 pensioners equally in each area.

The information was processed using mathematical calculations and statistical indicators and methods. For pair wise comparisons was used Fuller triangle concept, as well as the ordering method and mean values.

3. Results, Discussion

The aim of the research was to find out which biometric means respondents use most often in their communication. We investigated that they also used biometrics in which communication channels and networks. It is equally interesting for us to find out how and to what extent they consider biometrics to be able to protect identity, to what extent it would be interesting to find out where they could imagine their further use in practice and what ethical threats they consider and fundamental to privacy and identification.

The graph describes the practical use of different types of biometrics in individual groups. As we mentioned above, they all use a password. Likewise, not all groups use behavioural biometrics. Likewise, respondents did not encounter the term biochiping and do not use it.
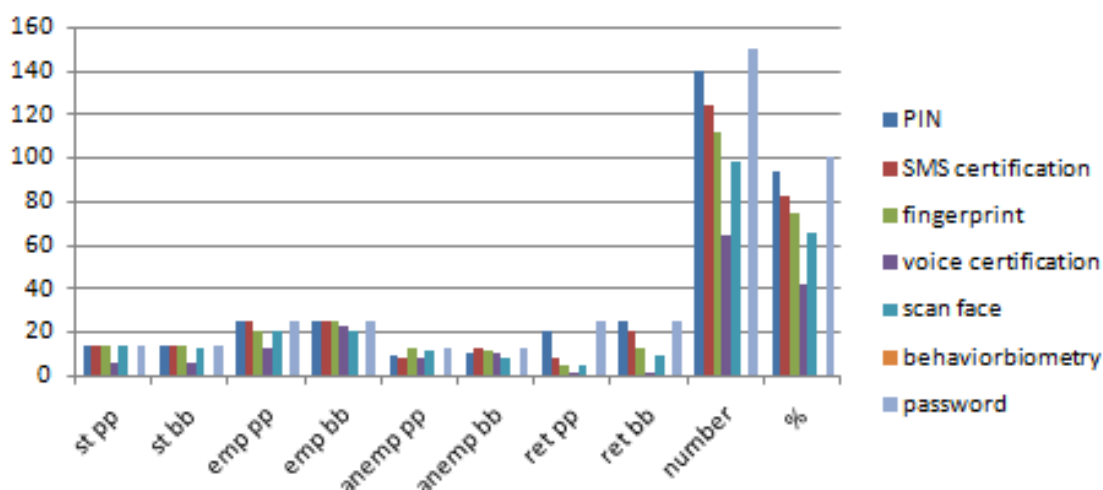
Figure 1: Using of different types of biometrics in individual groups

In the following graph we can see the most common places where respondents identify and log in with various forms of biometrics. The most common place is the household. Respondents meant this as a home WIFI network, as well as authentication via home alarms and inputs to IT equipment.

We see that a group of students use social networks more like Facebook (Pavlíček, 2008, 2010, 2014, 2020), which is the domain of workers. The least pensioners use social networks. Rather, their personal registration is associated with a practical overlap such as writing mail or working with the bank on a simple level. If we put it in context and interrelationships, we would find that education greatly affects the level of use of biometrics. The more educated the respondent, the more advanced the level of biometrics he uses. We also found out from relational research that the age of the respondents is not important. The students responded the most to the question in connection with biochiping, but at the same time they could not imagine its use in communication in social networks. They preferred personal login based on a password or biometrics. Already behaviour biometrics was a problem they did not want to accept.

	st pp	st bb	emp pp	emp bb	anemp pp	anemp bb	ret pp	ret bb
home	13	13	25	20	12	12	20	12
work occount	8	12	25	25	6	5	5	3
e mail users	13	13	25	25	12	10	20	25
bank, e-pay	13	13	25	25	12	12	18	12
twitter	10	13	12	8	0	2	3	5
instagram	13	10	15	20	8	5	5	5
fb	13	13	20	21	10	11	5	15

Figure 2: use of passwords according to groups of respondents and the type of service used

The final group of researches focused on the potenti: al threats associated with various biometric systems. The main threats were linked to concerns about restrictions on freedom and ethical abuse of biometrics. It turned out that there is a relationship - the link between knowledge (education) about biometric systems and their intensity of use. Voice biometrics is still acceptable, but behavioral is already a problem, and radical biochiping has been rejected.

If the behavioral biometrics of the respondents' answers were used in the examined areas, they would be most afraid of using them, because according to the answers they would most affect the natural way of human life. Paradox, the biggest problem for all respondents is the fear of fishing. This is mostly done through the use of an e-mail address, where the user is most at risk. The intensity of the password and its quality could be a source of further research.

The biggest form of fear affects pensioners and then the unemployed from the mining area. We can assume that for pensioners, it is associated with the fact that they do not understand the possible ways of protection and the principles of biometrics, and for the unemployed it is associated with education and access to information.

Biochiping from the results of the research, by the fact that we received negative answers in all cases, can not yet be talked about in Slovak society as a possible access to social networks. Respondents do not trust this technology for several reasons. The main reason is the threat to freedom, which was stated by up to 60% of respondents, the possibility of abuse by 30% and the same religious reason as stated by 50% of respondents in all groups and regions. It is in the religious context that the highest incidence is in all groups, so we can consider it as the main reason.

4. Conclusions

Biometrics has become part of everyone's daily life. It does not matter the region, its development, what we have examined, but the use and intensity of use of higher degrees of biometrics is associated with practical life, and with education. A high degree of relevance between education and the use of a higher degree of biometric protection has been demonstrated. Another interesting factor is economic activity where, there is a high degree of interaction between employment and retirement age and the perception of threat. Thus, the more economically active a person is, the higher the level of security through better biometrics gives the opportunity to get rid of the fear of danger, so it is clear that knowledge and understanding of the principles frees users.

Finally, we would like to mention that the use of biometrics and its better ways of verifying personality is associated more with practical life, such as communication at work and the banking sector. The use of biometrics in social networks is associated with the basic method (PIN, password), in a few cases (20%) with more complex methods such as voice recognition or face identification is used more basic than unlocking IT equipment or communicating with the bank.

We also found an interesting fact from the research that the middle age generation uses the most FB and the generation of students from another social network. We can therefore assume that it is more interesting for users - students to selectively influence the content of information, where individual users decide on the choice of individual social networks and their use (Instagram - images and movies, Twitter - communication), how to use global information social media such as Facebook.

5. References

BAČA, M., M. SCHATTEN a J. ŠEVA, 2009. Behavioral and physical biometric characteristics modeling used for its security improvement. In Transport problems [online]. Volume 4, Issue 4, pages 5-13 [cit. 2020-04-11]. ISSN (online edition) 2300-861X. In: http://transportproblems.polsl.pl/pl/Archiwum/2009/zeszyt4/2009t4z4_01.pdf

BÖHMOVÁ, L, NOVÁK, R. (2015). How Employers Use LinkedIn for Hiring Employees in Comparison with Job Boards. In: IDIMT-2015 Information Technology and Society Interaction and Interdependence. Linz: Trauner Verlag Universität, pp. 189–194. ISBN 978-3-99033-395-2.

ČERNÝ, J., POTANČOK, M., MOLNÁR, Z. (2019) Using open data and Google search data for competitive intelligence analysis. Journal of Intelligence Studies in Business JISIB [online]. 9/2, pp. 72–81. ISSN 2001-015X. available at: https://ojs.hh.se/index.php/JISIB/article/download/470/216.

CHIRCHI, V. R. E., L. M. WAGHMARE a E. R. CHIRCHI, 2011. Iris biometric recognition for person identification in security systems. In International Journal of Computer Applications [online]. Volume 24, number 9, pages 1-6 [cit.2020-04-11]. ISSN 0975 - 8887. In: http://www.ijcaonline.org/volume24/number9/pxc3874002.pdf

DOUCEK, P., PAVLÍČEK, L., SEDLÁČEK, J., NEDOMOVÁ, L. Adaptation of Password Strength Estimators to a Non-English Environment – the Czech Experience. Computers & Security [online]. 2020, roč. 95. ISSN 0167-4048. DOI: 10.1016/j.cose.2020.101757. Dostupné také z: https://www.sciencedirect.com/science/article/pii/S0167404820300419.

DRAHANSKÝ, M., F. ORSÁG. a all., 2011. Biometrie. Brno: Computer Press. ISBN 978-80-254-8979-6.

GASPAR , R., ROHDE, P., GIGER, J., Unconventional settings and uses of human enhancement technologies: A non-systematic review of public and experts' views on self-enhancement and DIY biology/biohacking risks [online]. In: Human behavior and emerging technologies. [cit.2020-04-11]. Volume 1, Issue 4, October 2019, https://doi.org/10.1002/hbe2.175

MLÁDKOVÁ, L., PROCHÁZKA, D. A. 2015 Leading Change in Management Education in Department of Management. In: ROUCO, José Carlos Dias (ed.). Proceedings of the 11th European Conference on Management Leadership and Governance. Lisabon, 12.11.2015 – 13.11.2015. Reading: Academic Conferences and Publishing Int. Ltd., pp. 248–253. ISBN 978-1-910810-76-7. ISSN 2048-9021.

PAVLICEK, A. (2010). Human Resources in ICT- Information and Technological Literacy as the Qualification Prerequisites. In P. Doucek, G. Chroust, & V. Oskrdal (Eds.), IDIMT-2010: INFORMATION TECHNOLOGY - HUMAN VALUES, INNOVATION AND ECONOMY (Vol. 32, pp. 153–159). Universitatsverlag Rudolf Trauner., ISBN: 978-3-85499-760-3 WOS:000288345500016

PAVLICEK, A. (2014). New and Social Media in Workplace. In IDIMT-2014: NETWORKING SOCIETIES - COOPERATION AND CONFLICT (pp. 135–144). Trauner Verlag. ISBN 978-3-99033-340-2. WOS:000346148300015

PAVLICEK, A., ROSICKY, A. (2008). Wiki System as a Knowledge Management Tool. In C. Gerhard, D. Petr, & K. Jan (Eds.), 16th Conference on Interdisciplinary Information Management Talks (Vol. 25, pp. 403–412). Universitatsverlag Rudolf Trauner. WOS:000265436800030

PAVLICEK, A., SYROVATKOVA, J. (2020) New Trends in Product Placement Strategies - Case of Instagram, Hradec Economic Days, Vol. 10(1), pp. 582 – 583, ISBN ISBN 978-80-7435-776-3, DOI: 10.36689/uhk/hed/2020-01-066

YOUNG-SIK J., JI S. P., JONG H. P., 2014. An efficient authentication system of smart device using multi factors in mobile cloud service architecture [online]. In International Journal of Communication Systems. [cit.2020-04-11]. Volume 27, Issue 4, 2014 https://doi.org/10.1002/dac.2694.

SMART SUPPLY CHAIN

SMART SUPPLY CHAIN: ARE THE PROCURERS READY FOR INDUSTRY 4.0

Michal Tkáč

The Faculty of Business Economics with seat in Košice,
The University of Economics in Bratislava
michal.tkac1@euke.sk

Stephen Kelly

Edge Hill University Business School
stephen.kelly@edgehill.ac.uk

Klaas Stek

University of Twente
Graz University of Technology
klaas.stek@utwente.nl

Abstract

Industry 4.0 brings changes to the supply chain environment and companies that want to benefit from these changes must be prepared to implement them. However, successful implementation requires skilled and trained personnel who will embrace these changes and know what to do. This paper identifies the skills and competencies required in Purchasing and Supply Management (PSM), which are often mentioned together with Industry 4.0. The purpose of the research is to determine how Industry 4.0 changes PSM requirements and which types of skills will be needed in the future. Based on a systematic literature review of Industry 4.0 oriented articles, that specifically mention PSM skills, we provide a current view on the importance of particular skill categories in the era of machine-to-machine communication.

1. Introduction

The European Parliament defines Industry 4.0 as "the comprehensive transformation of the whole sphere of industrial production through the merging of digital technology and the internet with conventional industry. In short, everything in and around a manufacturing operation (suppliers, the plant, distributors, even the product itself) is digitally connected, providing a highly integrated value chain." (Davies, 2015). A more focused definition of Industry 4.0 states that "Industry 4.0 is characterized by cyber-physical systems with autonomous machine-to-machine communication." (Schiele and Torn, 2020). Although a broadly accepted definition of Industry 4.0 is still lacking

(Brettel et al., 2014), the identification of future requirements for jobs and skills, which can ensure the successful implementation of Industry 4.0 is becoming an increasingly important task for researchers (Gehrke et al., 2015), policymakers (Smit et al., 2016) and corporate managers (Lorenz et al., 2015). The implementation of automation and digital technologies, such as Cyber-Physical Systems, the Internet of Things, cloud computing and Big Data in manufacturing, significantly affects work processes and the work environment (Doucek and Hološka, 2019 Chryssolouris et al., 2013). The content and nature of jobs is changing, and new kinds of jobs are being introduced. These changes also impact purchasing and supply management (PSM). Although there is an increase in the attention being paid to the changing role of PSM professionals, research on the changes in the jobs and skills required by Industry 4.0 is still emerging. Furthermore, the majority of Industry 4.0 and skills-oriented research (e.g., Gehrke et al., 2015; Hecklau et al., 2016) is mostly theoretical and tends to provide only a general overview of Industry 4.0 related competencies, without a clearly defined focused on the specific organizational areas where technologies are being implemented (Pinzone et al., 2017). Such approaches make it difficult to identify how specific job profiles should evolve and what types of skills will be required and needed in specific positions for Industry 4.0 to be effectively implemented (Dworschak & Zaiser, 2014). To mitigate these shortcomings in PSM, researchers from five universities formed an international project entitled PERSIST (Purchasing Education Research Syndicate: Industry 4.0 Skills), to analyse the role of PSM in the era of machine-to-machine communication (see: www.project-persist.eu). This paper presents an analysis based on the initial results of the project and provides an overview of PSM skills that have been identified in the Industry 4.0 oriented academic literature as being required for the modern PSM practitioner.

2. Methodology of the research

The research in this paper is based on a focused and systematic PSM Industry 4.0 oriented literature review. The literature review was done using the SCOPUS database and search was limited to scientific peer-reviewed articles, conference proceedings and book chapters published from 2012 to 2019. Because this review aims to identify procurement skills connected to Industry 4.0, the search string was divided into three segments:

Procurement segment: including synonyms and keywords connected with procurement, such as: PSM, supply chain, purchasing, procurement, buying, sourcing, contract management, supplier management, supply management, outsourcing, supplier-buyer relationship, inter-organisational relationship, buyer-supplier relationship, supplier selection, supplier evaluation, vendor selection, vendor evaluation.

Industry 4.0 segment: including synonyms and keywords connected with Industry 4.0. such as: industrial revolution, 4th, smart, IoT, IoE, digitisation, digitization, future, big data, automation, robotization, robotization, intelligent robotics, artificial intelligence, autonomous, Industry 4.0, lights out, machine dominant, cooperating machine, learning machine, interoperability, virtualisation, virtualization, virtual manufacturing, interconnect, cyber physical, cloud, machine to machine, smart factory, intelligent factory, real time factory, factory of things, data mining, ERP, augmented reality, digital industrial age, data processing, data economy, industrial internet of things, IIoT, CPS platform, networked manufacturing, Digital Manufacturing, data security, data analytics.

Skills segment: including synonyms and keywords connected with skills, such as: Competencies, training, education, skill, knowledge, competency, competence, digital literacy, task performance, self-efficacy, capability.

In the search strings the conjunction "AND" was used to connect different segments and "OR" was used to connect keywords within the segments. To ensure a focus on articles dealt with PSM and skills, the search was limited to capture words from the procurement segment and skills segment that

appeared in the title or abstract, or in the keywords of the article, whereas the keywords from industry 4.0 segment had to occur in the full text of the article. A preliminary search identified 3,651 articles, but a more detailed abstract analysis showed that only 49 articles were dealing with procurement skills. Distribution of reviewed articles based on the year of their publication is presented in Table 1 and the list of identified articles, together with their references are at the end of the article.

Table 1 Distribution of selected articles per year

	2012	2013	2014	2015	2016	2017	2018	2019	Total
Identified by search	335	381	436	419	482	460	580	558	3651
Selected articles	0	4	3	2	1	6	22	11	49

Source: Authors

The skills from these articles were then identified and classified based on the work of Bals et al. (2019), entitled "Purchasing and supply management (PSM) competencies: Current and future requirements", which is a recent work that developed an empirically based set of both current and future skills for PSM practitioners. This work followed the approach of Tassabehji and Moorhouse (2008) and some HRM studies "in order to obtain the widest possible view of competencies and enhance the understanding of what PSM professionals need to be able to do their jobs." (Bals et al. (2019: 2).

3. The research

To determine which PSM skills and how frequently a particular skill is mentioned in Industry 4.0 articles, the framework from Bals et al. (2019) was used to divide skills into four categories of Technical Skills, Interpersonal Skills, Internal/External Enterprise Skills, Strategic Business Skills. Based on the definition adapted from Tassabehji and Moorhouse (2008) the Technical skills represent "Fundamental and basic administrative skills necessary for any procurement professional in the 21st century" (Bals et al., 2019: 11). They include product knowledge, computer literacy, total quality management and government legislation. The sub-category "Advanced procurement process skills (APP)" includes category management, project planning and e-procurement (Bals et al., 2019: 11). Using the classification table mentioned in the methodology, the distribution of the technical skills within Industry 4.0 related academic literature is presented in Table 2.

Table 2 List of technical skills identified in the literature review

	Skills	2013	2014	2015	2016	2017	2018	2019	N
	Big Data Analytics	1	0	1	0	1	13	4	20
	Project Management	2	2	1	0	2	1	2	10
	Tools and Systems Implementation	2	0	1	0	0	3	1	7
	Negotiation	0	1	0	0	1	1	3	6
	Process optimization	0	1	1	0	0	3	1	6
	Database Management*	0	0	0	0	0	4	2	6
	Computer Literacy	0	0	1	0	0	2	2	5
	Basic knowledge on PSM role & processes	0	0	0	0	0	1	3	4
	Contract Management	0	1	0	0	0	1	2	4
	eProcurement Technology	0	0	1	0	0	0	3	4

	Skills	2013	2014	2015	2016	2017	2018	2019	N
Technical skills	Product knowledge	0	1	0	0	0	2	1	4
	Strategic sourcing	0	1	1	0	0	0	2	4
	Innovative sourcing approaches	0	0	0	0	0	1	2	3
	Innovation sourcing	0	0	0	1	1	0	1	3
	Statistical Skills*	0	0	1	0	0	0	2	3
	Cost savings	0	1	0	0	0	0	1	2
	Skill for Artificial intelligence*	0	0	0	0	0	0	2	2
	Operations research*	0	0	2	0	0	0	0	2
	Skills for Blockchain*	0	0	0	0	0	1	1	2
	Skills for Machine Learning*	0	0	0	0	0	1	1	2
	Skills for Distributed Computing*	0	0	0	0	0	2	0	2
	Intellectual Property	0	0	0	0	0	0	1	1
	KPI Reporting Design	0	0	0	0	0	0	1	1
	Languages	0	0	0	0	0	0	1	1
	Quality assurance	0	0	0	0	0	0	1	1
	Automation	0	0	0	0	0	0	1	1
	Mathematical modelling*	0	0	1	0	0	0	0	1
	Skills for Solution development*	0	0	0	0	0	1	0	1
	Skills for 3D Printing*	0	0	0	0	0	1	0	1
	Skills for Cloud computing*	0	0	0	0	1	0	0	1
*Skills not listed in Bals et al. (2019)									

Source: Authors

As can be seen in table 2, big data analytics is the most frequently mentioned PSM technical skill. This is followed by skills related to project management and tools and system implementation and the top five list of technical skills is concluded by negotiation and process optimisation skills. From the skills that were not listed in Bals et al. (2019), database management is mentioned quite often in the articles published more recently (i.e. after 2017).

Interpersonal skills represent the second group of identified skills, which are "necessary for interaction with people in teams and on an individual level including written and oral communication, conflict resolution, influencing and persuasion, group dynamics, leadership, problem solving and interpersonal and cultural awareness" (Tassabehji and Moorhouse, 2008; Bals et al., 2019). The list of interpersonal skills identified in the review are presented in Table 3.

Table 3 List of interpersonal skills identified in the literature review

	Skills	2013	2014	2015	2016	2017	2018	2019	N
Interpersonal skills	Analytical skills	2	1	1	0	2	4	3	13
	Decision making	0	1	1	0	0	4	3	9
	Teamwork-working in teams	2	1	1	1	1	0	3	9
	Interpersonal Communication	0	1	2	0	1	1	2	7
	Creativity	0	0	2	1	0	1	2	6

Skills	2013	2014	2015	2016	2017	2018	2019	N
Knowledge sharing	0	0	0	0	0	3	3	6
Leadership	0	1	1	0	1	1	2	6
Learning agility	1	0	0	1	0	1	2	5
Problem solving*	0	1	0	1	0	1	2	5
Conflict Resolution	0	1	2	0	0	0	1	4
Integrity	0	0	0	0	2	0	1	3
Results focus driving for results	0	0	0	1	1	0	1	3
Openness, Open-minded,	0	0	0	1	0	0	2	3
Passion	0	0	1	0	1	0	1	3
Effective questioning techniques	0	0	1	0	0	0	1	2
Curiosity	0	0	0	1	0	0	1	2
Self-confidence	0	0	0	0	1	0	1	2
Prioritization	0	0	0	0	0	0	1	1
Remote Virtual Working	0	0	0	0	0	0	1	1
Structured way of working	0	0	0	0	0	0	1	1
Deal with ambiguity	0	0	0	0	0	0	1	1
Humility	0	0	0	0	0	0	1	1
Mobility	0	0	0	0	0	0	1	1
Resilience	0	0	0	0	0	0	1	1
Self-reflection	0	0	0	0	0	0	1	1
Self-reliance	0	0	0	0	0	0	1	1
Flexible thinking*	0	0	0	1	0	0	0	1
Stress Management*	0	0	1	0	0	0	0	1
Awareness of the needs of other*	0	0	1	0	0	0	0	1

*Skills not listed in Bals et al. (2019)

Source: Authors

In terms of interpersonal skills, the results showed that analytical skills are the most mentioned in the Industry 4.0 oriented scientific literature. This is followed by decision making (2nd) and team working (3rd), interpersonal communication (4th) and creativity (5th). In addition to the skills listed in Bals et al. (2019), problem-solving is the only skill mentioned more than once.

The third group of skills are Internal/External enterprise skills, which, from an internal perspective, are those that "relate to the overall business and how the different functions interact" (Tassabehji and Moorhouse, 2008; Bals et al., 2019), while external enterprise skills are related "to the supply chain/network and its stakeholders". (Tassabehji and Moorhouse, 2008; Bals et al., 2019). The skills identified within this category can be found in Table 4.

Table 4 List of Internal/External enterprise skills identified in the literature review

	Skills	2013	2014	2015	2016	2017	2018	2019	N
Internal/External enterprise skills	Cross-functional abilities & knowledge	2	0	0	1	0	5	2	10
	Logistics	2	1	1	0	1	3	1	9
	Communication skills	2	0	1	0	1	1	3	8
	Change Management	2	1	0	0	0	0	3	6
	Supply Chain	2	0	2	0	0	1	1	6
	Supplier management	0	2	0	0	1	1	2	6
	Stakeholder Relationship Management	1	1	0	0	0	1	2	5
	Finance	0	0	1	0	0	1	2	4
	Marketing	2	0	1	0	0	0	1	4
	Customer Focus	0	1	0	0	1	0	2	4
	Manufacturing/Production	0	1	0	0	1	0	1	3
	Cultural awareness	0	0	0	0	0	0	3	3
	Business ethics*	0	0	1	0	0	1	1	3
	International regulations/Legal knowledge*	0	0	1	0	0	1	1	3
	Engineering	0	1	0	0	0	0	1	2
	Quality (QHSE)	0	0	1	0	0	0	1	2
	R&D	0	1	0	0	0	0	1	2
	Sales	0	0	0	0	0	1	1	2
	Operations management*	0	1	1	0	0	0	0	2
	Networking	0	0	0	0	0	0	1	1
	Customer relationship management*	0	0	1	0	0	0	0	1
	Warehousing and distribution*	0	0	1	0	0	0	0	1
	Warehouse management*	0	0	1	0	0	0	0	1
	Transportation management*	0	0	1	0	0	0	0	1
*Skills not listed in Bals et al. (2019)									

Source: Authors

From the Internal/External enterprise skills identified by literature review, cross-functional abilities & knowledge are those mentioned most often in the Industry 4.0 oriented scientific literature. This is followed by logistic skills (2nd), communication skills (3rd), knowledge of change management (4th) and supply chain (5th). From the skills not listed in Bals et al. (2019), Business ethics and International regulations/Legal knowledge are mentioned more than once.

The final category is Strategic business skill, which: "…relate to broader strategic issues and how procurement can impact on overall organisational value such as planning and managing strategic partnerships and alliances, risk management and adding value to the organisation" (Tassabehji and Moorhouse, 2008; Bals et al., 2019) and these skills identified in the literature review are shown in Table 5.

Table 5 List of Strategic business skills identified in the literature review

	Skills	2013	2014	2015	2016	2017	2018	2019	N
Strategic business skills	Strategic thinking	2	1	1	0	0	3	5	12
	Business Acumen	1	1	2	0	1	3	3	11
	Holistic Supply Chain Thinking	2	0	0	0	1	5	1	9
	Risk management	2	1	1	0	2	0	2	8
	Financial acumen	3	0	1	0	0	1	1	6
	Sustainability	2	0	0	0	0	0	2	4
	Future thinking/Innovativness*	0	0	0	0	0	4	0	4
	Critical thinking	0	0	0	0	1	0	2	3
	PSM Best Practice Intelligence Scouting	0	1	0	0	0	1	1	3
	Planning*	1	0	1	0	0	1	0	3
	Technological trends*	0	0	0	0	0	1	0	1
*Skills not listed in Bals et al. (2019)									

Source: Authors

As shown in Table 5, strategic thinking represents the most frequently mentioned Strategic business skill, with Business acumen in second and holistic supply chain thinking (3rd), risk management (4th) and financial acumen (5th). From the skills not listed in Bals et al. (2019), Future thinking/Innovativeness and planning are mentioned more than once.

Table 6 below synthesises the results at an overall category level and shows how frequently each group of procurement skills are mentioned in the Industry 4.0 oriented scientific literature.

Table 6 Distribution of skills categories

Skills	2013	2014	2015	2016	2017	2018	2019	Total	Top 10
Technical Skills	5	8	11	1	6	38	41	111	2
Interpersonal Skills	5	7	14	8	10	16	40	101	3
Internal/External Enterprise Skills	13	10	14	1	5	16	30	90	2
Strategic Business Skills	13	4	6	0	5	19	17	68	3

Source: Authors

In terms of frequency, Table 6 shows that technical skills are the most represented skill group, although only two technical skills occur in the Top 10 list of most frequently mentioned skills. However, it is worth mentioning that in Industry 4.0 oriented articles, the Interpersonal skills group is the second largest group and three skills from this group are in the Top 10 list. Interestingly, the occurrence of strategic business skills was the smallest within the researched literature, yet three such skills are in the TOP 10 list.

To compare how PSM skillset determined by Industry 4.0 oriented articles differs from a general PSM one, we compared our results with the list of PSM competencies defined in Bals et al. (2019). In 2016, they conducted a series of interviews with PSM professionals and formulated a list of Top 10 current PSM competencies and a list of Top 10 future PSM competencies. The comparison of our Top 10 Industry 4.0 oriented PSM competencies with their lists is presented in Table 7.

Table 7 Comparison of Top 10 current PSM competencies, Future PSM competencies and Industry oriented 4.0 PSM competencies

Current PSM Competencies	Future PSM Competencies	Industry 4.0 oriented PSM Competencies
Analytical skills	Analytical Skills	Analytical skills
Strategic thinking	Strategic thinking	Strategic thinking
Basic knowledge on PSM role & processes	Big Data Analytics	Big Data Analytics
Communication skills	Holistic supply chain thinking	Holistic Supply Chain Thinking
Cross-functional abilities & knowledge	Automation	Cross-functional abilities & knowledge
Interpersonal Communication	Computer Literacy	Business Acumen
Negotiation	eProcurement Technology	Decision making
Stakeholder Relationship Management	Process optimisation	Logistics
Strategic sourcing	Strategic Sourcing	Project Management
Sustainability	Sustainability	Teamwork-working in teams

Source: Modified by authors based on Bals et al. (2019)

The comparison shows some areas where Industry 4.0 PSM oriented competencies align with the more general current and future PSM competencies. Analytical skills and Strategic thinking represent two types of skills where there is an alignment across all findings. Big Data Analytics and Holistic Supply Chain thinking connect Industry 4.0 PSM oriented competencies with Future PSM competencies and Cross-functional abilities & knowledge match Current PSM competencies with the Industry 4.0 oriented PSM competencies.

4. Conclusion

Scientific literature dealing with the implementation of Industry 4.0. or its equivalents in the broad field of supply chain management is growing. Most of the articles are introducing or evaluating new technologies, models and approaches to increase quality, performance, or reduce risk in the supply chain as a whole. Although implementation of all these enhancements is usually described in detail in a particular article, specific details of the skills and competencies required for their use are usually missing. To contribute to this emerging field, this systematic literature review analysed all articles which deal with PSM skills in Industry 4.0 oriented literature. From 3,651 identified articles, 49 of them were recognised as Industry 4.0 oriented and mentioned concrete skills linked to PSM. A detailed analysis of these articles showed that nine years after its introduction, Industry 4.0 does not seem to have changed the PSM skillset in a significant way. The majority of the PSM skills mentioned in Industry 4.0 oriented articles can be defined as traditional PSM skills, which are frequently mentioned in studies over the last 20 years. Apart from Big data analytics, typical Industry 4.0 oriented skills like Blockchain skills, Machine Learning skills, Distributed Computing skills, 3D Printing skills, Cloud computing skills, are only mentioned occasionally. Moreover, although Industry 4.0 should be more autonomous and machine-to-machine oriented, interpersonal skills are frequently mentioned in PSM Industry 4.0 oriented literature. On the other hand, with the growth of Industry 4.0, technical skills are becoming increasingly important and appear to supersede those skills of a more interpersonal nature. Further analysis also showed that, in articles which mention concrete Industry 4.0 purchasing technologies, technical skills such as analytical skills and big data analytics

are becoming more significant and, together with strategic thinking and taking a supply chain holistic view, represent the most required skills for PSM practitioners in the future Industry 4.0 environment. It should be noted that from 49 reviewed articles only 35 articles were able to link PSM with concrete Industry 4.0. purchasing technology. Furthermore, the vast majority of such articles do not describe how to train and teach such PSM Industry 4.0 skills. Based on these results, it is suggested that PSM environments have not yet engaged with the full potential of the technological advances behind Industry 4.0, which may leave PSM practitioners unprepared for further changes.

5. Acknowledgement

6. References

Arunachalam, D., Kumar, N., & Kawalek, J. P. (2018). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. Transportation Research Part E: Logistics and Transportation Review, 114, 416-436.

Asdecker, B., & Felch, V. (2018). Development of an Industry 4.0 maturity model for the delivery process in supply chains. Journal of Modelling in Management.

Bake, J. S., Pessôa, M. V. P., & Becker, J. M. J. (2018). Service Chain Logistics Management for Increasing Equipment Uptime. Procedia CIRP, 73(1), 210-215.

Bals, L., Schulze, H., Kelly, S. & Stek, K. (2019). Purchasing and supply management (PSM) competencies: Current and future requirements. Journal of Purchasing and Supply Management. Vol 25 (5).

Bausman, D., Chowdhury, M., & Tupper, L. (2014). Best practices for procurement and management of professional services contracts. Journal of Professional Issues in Engineering Education and Practice, 140(3), 04013019.

Biazzin, C., & Carvalho, L. C. (2019). BIG DATA IN PROCUREMENT: THE ROLE OF PEOPLE BEHAVIOUR AND ORGANIZATION ALIGNMENT. Dimensión Empresarial, 17(4).

Cacciolatti, L., Lee, S. H., & Molinero, C. M. (2017). Clashing institutional interests in skills between government and industry: An analysis of demand for technical and soft skills of graduates in the UK. Technological Forecasting and Social Change, 119, 139-153.

Choi, Y., Lee, H., & Irani, Z. (2018). Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector. Annals of Operations Research, 270(1-2), 75-104.

Chopra, A. (2019). AI in Supply & Procurement. In 2019 Amity International Conference on Artificial Intelligence (AICAI) (pp. 308-316). IEEE.

Daghfous, A., & Belkhodja, O. (2019). Managing Talent Loss in the Procurement Function: Insights from the Hospitality Industry. Sustainability, 11(23), 6800.

Daxböck, C., Kröber, J., & Bergmann, M. (2019). Digitized Performance Management Along the Supply Chain. In Performance Management in Retail and the Consumer Goods Industry (pp. 405-423). Springer, Cham.

De Vass, T., Shee, H., & Miah, S. J. (2018). The effect of "Internet of Things" on supply chain integration and performance: An organisational capability perspective. Australasian Journal of Information Systems, 22.

Ding, B. (2018). Pharma industry 4.0: Literature review and research opportunities in sustainable pharmaceutical supply chains. Process Safety and Environmental Protection, 119, 115-130.

Drakaki, M., Gören, H. G., & Tzionas, P. (2018). An intelligent multi-agent based decision support system for refugee settlement siting. International journal of disaster risk reduction, 31, 576-588.

Dubey, R., & Gunasekaran, A. (2015). Supply chain talent: the missing link in supply chain strategy. Industrial and Commercial Training.

Dubey, R., Gunasekaran, A., Childe, S. J., Luo, Z., Wamba, S. F., Roubaud, D., & Foropon, C. (2018). Examining the role of big data and predictive analytics on collaborative performance in context to sustainable consumption and production behaviour. Journal of Cleaner Production, 196, 1508-1521.

Eltantawy, R. A., & Giunipero, L. (2013). An empirical examination of strategic sourcing dominant logic: Strategic sourcing centricity. Journal of Purchasing and Supply Management, 19(4), 215-226.

Ha-Brookshire, J., McAndrews, L., Kim, J., Freeman, C., Jin, B., Norum, P., ... & Marcketti, S. (2017). Moral education for sustainable development: Exploring morally challenging business situations within the global supply chain context. Sustainability, 9(9), 1641.

Handfield, R., Jeong, S., & Choi, T. (2019). Emerging procurement technology: data analytics and cognitive analytics. International Journal of Physical Distribution & Logistics Management.

Hawkins, T. G., & Muir, W. A. (2014). An exploration of knowledge-based factors affecting procurement compliance. Journal of public procurement, 14(1), 1.

Hofmann, E., & Rutschmann, E. (2018). Big data analytics and demand forecasting in supply chains: a conceptual analysis. The international journal of logistics management.

Hofmann, E., Sternberg, H., Chen, H., Pflaum, A., & Prockl, G. (2019). Supply chain management and Industry 4.0: conducting research in the digital age. International Journal of Physical Distribution & Logistics Management.

Hopkins, J., & Sohal, A. (2019). An investigation into talent shortages in the Australian procurement profession. Higher Education, Skills and Work-Based Learning.

Jeble, S., Dubey, R., Childe, S. J., Papadopoulos, T., Roubaud, D., & Prakash, A. (2018). Impact of big data and predictive analytics capability on supply chain sustainability. The International Journal of Logistics Management.

Knight, L., Tu, Y. H., & Preston, J. (2014). Integrating skills profiling and purchasing portfolio management: An opportunity for building purchasing capability. International Journal of Production Economics, 147, 271-283.

Kosmol, T., Reimann, F., & Kaufmann, L. (2019). You'll never walk alone: Why we need a supply chain practice view on digital procurement. Journal of Purchasing and Supply Management, 25(4), 100553.

Loukis, E., Arvanitis, S., & Kyriakou, N. (2017). An empirical investigation of the effects of firm characteristics on the propensity to adopt cloud computing. Information Systems and e-Business Management, 15(4), 963-988.

Mandal, S. (2018). An examination of the importance of big data analytics in supply chain agility development. Management Research Review.

Mandal, S. (2019). The influence of big data analytics management capabilities on supply chain preparedness, alertness and agility. Information Technology & People.

Mazzola, L., Waibel, P., Kaphanke, P., & Klusch, M. (2018). Smart process optimization and adaptive execution with semantic services in cloud manufacturing. Information, 9(11), 279.

Mishra, A., Sinha, K. K., & Thirumalai, S. (2017). Project quality: The Achilles heel of offshore technology projects?. IEEE Transactions on Engineering Management, 64(3), 272-286.

Mohelska, H., & Sokolova, M. (2018). Management approaches for Industry 4.0–the organizational culture perspective. Technological and Economic Development of Economy, 24(6), 2225-2240.

Mora, A., Riera, D., Gonzalez, C., Arnedo-Moreno, J. (2015). A Literature Review of Gamification Design Frameworks. VS-Games 2015 - 7th International Conference on Games and Virtual Worlds for Serious Applications, art. no. 7295760.

Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. Information Systems Frontiers, 20(2), 209-222.

Prajogo, D., & Sohal, A. (2013). Supply chain professionals: A study of competencies, use of technologies, and future challenges. International Journal of Operations & Production Management, 33(11-12), 1532-1554

Qiao, R., Zhu, S., Wang, Q., & Qin, J. (2018). Optimization of dynamic data traceability mechanism in Internet of Things based on consortium blockchain. International Journal of Distributed Sensor Networks, 14(12), 1550147718819072.

Queiroz, M. M., & Telles, R. (2018). Big data analytics in supply chain and logistics: an empirical approach. The International Journal of Logistics Management.

Rajakaruna, S., Wijeratne, A. W., Mann, T. S., & Yan, C. (2017). Identifying key skill sets in humanitarian logistics: Developing a model for Sri Lanka. International journal of disaster risk reduction, 24, 58-65.

Rajkumar, C., & Stentoft, J. (2017). Harnessing capabilities and practices for sourcing innovation: an exploratory study. Logistics Research, 10(10), 1-19.

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. Journal of Business Logistics, 36(1), 120-132.

Sergeeva, S. A. (2019). Monitoring of Educational Programs for the Development of Professional Competencies by Procurement Specialists in Procurement Statistics. Academic Journal of Interdisciplinary Studies, 8(2), 239-239.

Shafique, M. N., Rashid, A., Bajwa, I. S., Kazmi, R., Khurshid, M. M., & Tahir, W. A. (2018). Effect of IoT capabilities and energy consumption behavior on green supply chain integration. Applied Sciences, 8(12), 2481.

Shamim, S., Cang, S., Yu, H., & Li, Y. (2016, July). Management approaches for Industry 4.0: A human resource management perspective. In 2016 IEEE Congress on Evolutionary Computation (CEC) (pp. 5309-5316). IEEE.

Sun, L., & Zhao, L. (2018). Technology disruptions: exploring the changing roles of designers, makers, and users in the fashion industry. International Journal of Fashion Design, Technology and Education, 11(3), 362-374.

van Hoek, R., Wagner, B., & Sohal, A. S. (2013). Developing competencies of supply chain professionals in Australia: collaboration between businesses, universities and industry associations. Supply Chain Management: An International Journal.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics, 34(2), 77-84.

Williams, A. M., Lau, F., & McCue, C. P. (2018). Acknowledging knowledge: The perception of knowledge requirements for public procurement officials and their professional development. Journal of Public Procurement, 18(1), 50-67.

Yu, W., Chavez, R., Jacobs, M. A., & Feng, M. (2018). Data-driven supply chain capabilities and performance: A resource-based view. Transportation Research Part E: Logistics and Transportation Review, 114, 371-385.

Zhou, L., Wang, L., Ai, T., & Sun, Y. (2018). BeeKeeper 2.0: confidential blockchain-enabled IoT system with fully homomorphic computation. Sensors, 18(11), 3785.

Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective. International journal of mechanical, industrial science and engineering, 8(1), 37-44.

Chryssolouris, G., Mavrikios, D., & Mourtzis, D. (2013). Manufacturing systems: skills & competencies for the future. Procedia CIRp, 7(2013), 17-24.

Davies, R. (2015). Industry 4.0 Digitalisation for productivity and growth. European Parliamentary Research Service, 10.

Doucek, P., Hološka, J (2019). Digital economy and industry 4.0 IDIMT-2019. Innovation and Transformation in a Digital World. 27th Interdisciplinary Information Management Talks (Sept. 4–6, 2019, Kutná Hora, Czech Republic). TRAUNER Druck GmbH & Co KG, Linz, 33-39

Dworschak, B., & Zaiser, H. (2014). Competences for cyber-physical systems in manufacturing–first findings and scenarios. In Procedia CIRP (Vol. 25, No. C, pp. 345-350). Elsevier.

Gehrke, L., Kühn, A. T., Rule, D., Moore, P., Bellmann, C., Siemes, S., & Standley, M. (2015). A discussion of qualifications and skills in the factory of the future: a German and American perspective. VDI/ASME Industry, 4, 1-28.

Hecklau, F., Galeitzke, M., Flachs, S., & Kohl, H. (2016). Holistic approach for human resource management in Industry 4.0. Procedia Cirp, 54(1), 1-6.

Lorenz, M., Rüßmann, M., Strack, R., Lueth, K. L., & Bolle, M. (2015). Man and machine in industry 4.0: How will technology transform the industrial workforce through 2025. The Boston Consulting Group, 2.

Pinzone, M., Fantini, P., Perini, S., Garavaglia, S., Taisch, M., & Miragliotta, G. (2017, September). Jobs and skills in Industry 4.0: An exploratory research. In IFIP International Conference on Advances in Production Management Systems (pp. 282-288). Springer, Cham.

Schiele, H. & Torn, R.J. (2020) Cyber-physical systems with autonomous machine-to-machine communication: Industry 4.0 and its particular potential for purchasing and supply Management Int. J. Procurement Management,

Smit, J., Kreutzer, S., Moeller, C., & Carlberg, M. (2016). Industry 4.0: Study. European Parliament.

Tassabehji, R., & Moorhouse, A. (2008). The changing role of procurement: Developing professional effectiveness. Journal of purchasing and supply management, 14(1), 55-68.

SIMULATION BASED BUSINESS MODELLING FOR GEOSPATIAL DATA SERVICE

Radoslav Delina, Marek Macik, Renata Olejarova

Technical University of Kosice, Faculty of Economics, Kosice, Slovakia
radoslav.delina@tuke.sk, marek.macik@tuke.sk, renata.olejarova@tuke.sk

Keywords

Business modelling, geospatial data, satellite data, simulation

Abstract

Geospatial data services are emerging on the market with new value for end users. The suitable business model is crucial to understand real value for end users and investors. Simulation based approach providing possibility to considering sensitivity of satellite image resolution on data service efficiency in line with providing simulated scenarios for optimizing the user acceptance can be promising enrichment of current approaches to business modelling. The paper provides the framework for this approach and potential enhancement on the example of data service for undeclared building diagnostics. On the base of simulation SW, the model and outputs are presented.

1. Introduction

The development of new technologies and principles has an increasing impact on the daily lives of people around the world. Whether it is process automation, increasing the efficiency of industry and technology, the development of modern procedures in industry, agriculture, but also in data analytics, financial markets, research and especially in the IT sector (MIT Libraries, 2019). In simplicity, we can say that technology enormously affects the daily lives of people around the world, and their impact is increasing and increasing (Tolga et al, 2013).

These technologies include the use of geospatial data, which has a real use in most industries and has made life easier for most people for more than a decade. Geospatial data can be known to the greatest extent from digital maps such as google maps, mapy.cz and various other map and infrastructure systems. However, geospatial data itself has a much wider use both in industry and in the simplification of many processes and services for residents, companies and the state itself (Castelein, et al, 2010; Dasgrupta, A. 2013)). Likewise, this data can be used extensively to prevent global disasters, speed up decision-making processes, improve the environment and reduce the environmental impact of society as a whole.

Geospatial or simplified just spatial data is topic that becomes interesting in past years for individuals, companies and even states and global organizations. Use of spatial data can be new way how to increase effectivity of products, services and even entire global economies (Krek, A. 2005).

Interest in special data and their impact on economy is subject of many international scientific studies as well as political steps in legislation. Development of digitalization and IT supports the idea of use big data where spatial data are part of them. With more common use of big data it is much easier to use and work with spatial data where companies and states are aware of it and use these technologies and data to make better innovation and way how to improve economic environment and society.

2. Theoretical background to business modelling

The concept of business model has come to the attention of academics and practitioners relatively recently, although it is perceived by society only for a relatively short time and its origins date back to the mid-20th century. The term has appeared in the literature for more than half a century, but there is still no uniform definition of the term, and these definitions are often very different and distant. The beginnings of the definition of the term business model were defined by Peter Drucker (1994), who considered the business model to be something that answers the questions: Who is the customer? What does the customer consider valuable and useful? How do you make money in a relevant business? What is the economic logic that explains how to bring value to customers at a reasonable cost? The definitions of some economists answer these questions directly or indirectly. For example, Alan Afuah (2004) considers a business model to be a money-making system that includes a set of activities that a company performs in a certain way and at a certain time in order to offer its customers the desired benefits while making a profit. It further defines this idea as a set of activities that an enterprise carries out in a certain way and at a certain time, using its resources to carry out activities that are common in the industry and thus creates value for the customer. Joan Magretta (2002) similarly defines the business model to be essentially a story that explains how a business works. Similarly, it defines Teece (2010), who considers the business model as a way for a company to deliver value to customers and turn payments received into profit. David Hunger and Thomas Wheelen (2008) see the business model as a method that the company uses to make money in the business environment, respectively the way it generates revenue and generates profit. Zott and Amit (2010) approach the business model somewhat differently, considering the business model as a system of interdependent activities that transcend the central enterprise and transcend its boundaries. They say that a separate business model and its activities are organized around a central company that maintains relationships with partners in a certain area and creates value together with partners, where it appropriates a share of the value created. Roman Casadesus-Masanell and Joan E. Ricard (2011) view the business model significantly differently. They consider such decisions to be the method of remuneration, procurement, location of operations, the extent of vertical integration, sales, marketing and others. These decisions have consequences that the authors consider to be the second component of the business model. The choice of price affects the volume of sales, which affects economies of scale and bargaining power. As a result, the company's ability to create value and its realization is affected. Simply put, they try to say that the business model consists of a set of managerial decisions and their consequences.

We would find dozens of such different definitions in the literature, where each definition would be different but in almost all we could notice that it contains three basic elements. These elements are products or services, the way the company is organized and the way how to profit. Otherwise, we could say that it is about what the company offers to customers, how it delivers these products or services and how it produces value from them - profitable model. Almost all definitions contain these 3 elements, but they differ in the emphasis on individual elements. The business model itself can be limited when it focuses only on the company or open when it also includes partners Hunger, (D. L., & Wheelen, T. L. 2008). The economic concept of the model emphasizes the economics of the company as a closer look. The systemic or conceptual approach emphasizes more complex, better said the full functioning of the business including the resources, processes and benefits for the customer, therefore a broader view of the business model.

3. Methodology

The purpose of this paper is to present more suitable business model approach based on risk and scenario simulation as current approaches. The main benefit is to adapt specific character of geospatial data to commercial data service understanding esp. in the context of adapting the most suitable balance between image resolution, price and users' acceptance or utility. The logic behind is based also on data service price flexibility from user perspective, where for example some services like tax inefficiency calculation from undeclared or shade properties/buildings. The amount of tax inefficiency and willingness to pay price for the data service will be depended on the price of geospatial data. The model should be able to optimize the value of the data service where right image resolution and related price will be selected based on predicted undeclared properties/buildings, value of related tax evasion and price of the satellite data. That's why and based on the examined approaches, we proposed a concept based on stochastic modelling. We have developed a model for simulating alternatives of performance characteristics that takes into account uncertainty and risks at different cost and revenue levels. This paper show only example how this new approach for business modelling should be applied and performed. The further research for modelling more specific features of satellite images and our research enhancement would be welcome.

For the analysis of business models, TCO (Total Cost of Ownership) approaches extended by identified properties of geospatial data was considered. The cost calculation is then determined by the flexibility in the selection of parameters and satellite data collection options. These multi-parameter characteristics and specifics for the identification and stochastic calculation of TCO and TBO (Total Benefit Opportunities) should be identified within following determinants (Purchasing & Procurement Center. 2020):

- different types of data (financial, sociodemographic, behavioural, sensory, SDI / Geo…)
- methods of their collection (aerial, satellite, IoT, terrestrial / frequency, experimental, existing databases eg socio - demographic and economic data, etc.)
- methods of their initial processing (statistics, visualization, modelling and simulation, including virtual reality, image processing, extensive data / data mining, data vs information vs knowledge…)
- possibilities of their sharing and synergies (additional value of aggregation and data sharing, possibilities of aggregation based on standardization of data models, motivational modelling, experimental verification and deployment of motivational models for data sharing…)
- possibilities of their practical use (domain, multisectoral use, support of data APIs ...)
- possibilities of their analysis and evaluation
- options, system, and frequency of their updates
- possibilities of their valuation and business models (data valuation methods, negotiation of conditions, contracting in the field of data use, new forms of business platforms and business models ...)
- problems in their use (legislative, personal / sensitive, missing standards, security, anonymity vs openness ...)
- design and development of negotiating, contracting and trustworthy tools for contracting data according to their specific types, properties and limitations
- design and development of API for easy interconnection of developed and existing SW applications with the potential of using the proposed data solution

4. Stochastic and simulation based data service business modelling approach

The example of this simulation based business model is focused on undeclared buildings diagnostic data service. When creating a simulation model, we have identified several basic input variables that take into account generic cost items, e.g. data quality, speed and timeliness of collection, need for update, type of collected data, etc. These variables were assigned calculation methods which in some cases took into account uncertainty and therefore had to be quantified from a risk point of view by establishing an appropriate statistical distribution. For example, prices for data in a given resolution (data quality) were collected from various sources that offered a range of prices and price ranges. Based on different values, we estimated the risk through a statistical distribution used for financial model simulation.

For simulation, we decided to use the software environment Palisade Corp., specifically Decision Tools Suite. One of the input variables was the identification of the resolution of satellite images, which were enriched with statistical distributions defining a certain coefficient of error or accuracy and at the same time suitability for a given type of data product. As example in table 1, is the creation of TCO for data services for the identification of undeclared properties, where the usability levels of individual resolutions have been defined as well as their error rate. Based on these risk quantifications, these fields were further used to model the business model setup

Table 1 Creation of TCO for data services

Scale		Resolution
1	Great	0,95
2	Very good	0,816666667
3	Good	0,595
4	Applicable with restrictions	NA
5	Not applicable	NA

Error rate vs need for minimal resolution	Area
15-30	0,895358565
30	0,816666667
40	0,7
50	0,595
55	Not applicable
60	Not applicable

The introductory element of the model was the table for determining the generic TCO, where it is possible to select the type of spectral parameters as well as geometric ones. According to these characteristics suitable for a given type of service, the model was created so the user defines the required minimal or exact resolution. In case of minimal required resolution, the model was constructed to simulate the entire business model for all resolutions equal to or better defined. According to these values, the corresponding error value was then loaded into the calculations according to the simulated value of the quantified statistical distribution.

Table 2 Table of input variables into TCO model

TCO

Spectral parameters	RGB			Risk		Simulation value
Geometrical parameters	*(RGB, hyperspectral, termo, multispectral, PAN, PANS, radar, shortwave infra)*					
Data quality			*ice level*			50
Minimal demanded resolution				3	50	
					0,595	
Exact demanded resoultion						
Error rate					0,595	
Demanded actualisation		yearly			1	
Type of images		satelite				
Update requirements				Clients portfolio plan		Price per km2
archive 90+ days	90+	RGBStereo90+			0,95	35
archive max. 90 days	90	RGBStereo90			0	55
priority imagery	P	RGBStereoP		15	0,05	99
standard imagery	S	RGBStereoG		11	0	55

In the next step, the model requires defining the necessary timeliness of data, which in a given product corresponds to four basic variants according to the type of spectrum and geometric parameters. Unique combinations of these characteristics are modeled and quantified by price, and which are inserted on the basis of a link to the table of identified prices per km^2 as seen in table 3.

Table 3 Cost of different data types and resolutions (€)

		archive 90+ days				archive max. 90 days				standard imagery				priority imagery			
		MONO		Stereo		MONO		Stereo		MONO		Stereo		MONO		Stereo	
		RGB	PAN	RGB	PAN	RGB	PAN	RGB	PAN	RGB	PAN	RGB	PAN	RGB	PAN	RGB	PAN
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
30	1	22,5	19	45	38	32,5	29	65	58	32,5	29	65	58	54,5	51	109	102
40	2	19,5	16	39	32	29,5	26	59	52	29,5	26	59	52	51,5	48	103	96
50	3	17,5	14	35	28	27,5	24	55	48	27,5	24	55	48	49,5	46	99	92
55	4	8	8	16	16	8	8	16	16	18	18	36	36	28	28	56	56
70	5	8	6	16	16	8	8	16	16	16	15	32	32	26	21	52	52
75	6	7	7	12	12	7	7	12	12	14	14	24	24	26	26	42	42
80	7	8	4,5	20	20	10	4,5	X	X	15	4,5	X	X	27	X	X	X
100	8	4	4	8	8	4	4	8	8	8	8	16	16	13,5	12	27	27

After quantifying these variables and other variables on the cost side entering the TCO model as standard, the model was created by determining business revenue models according to individual types of end customers, their specific prices and in terms of potential simulation of all possible decision scenarios. This approach makes it possible to retrospectively analyze backward scenarios in the model according to the required changes of the resulting values. It can be seen in table 4

Table 4 Individual types of end customers

Planned clients	Category		Planned range of mapping			Price for range
number of clients - organizations	1	km2	15		1	525
number of clients - public	2	km2	11,66666667		1	641,6666667
number of clients - government	3					
number of clients - towns, cities	4	km2	638,3333333	0,95	0,05	24384,33333

After defining and testing of all statistical distributions based on input source data, the simulation model was set to the Latin Hypercube method, which provides relatively accurate results across the width of the simulated statistical distribution interval despite the lower number of recalculated iterations, which we set to 10,000. Using SimTable function model defined all levels of the specified resolution (50, 40, 30cm) and prepared the model for simulations and recalculations of the whole model for each level of resolution separately and for the purposes of effective comparison of scenarios.

The result of the simulation was a comparison of gross profit, which was the main interest of the investor, in all three scenarios according to the possible distinctions of the acquired data. The following table 5 shows a comparison of the profitability density histograms of the three scenarios, where the claret color corresponds to the 50cm, blue 40cm and green 30cm resolution of the images used together with the cumulative frequency, which clearly defines all three scenarios according to risk-return.

Table 5 Comparison of the profitability density histogram of three scenarios
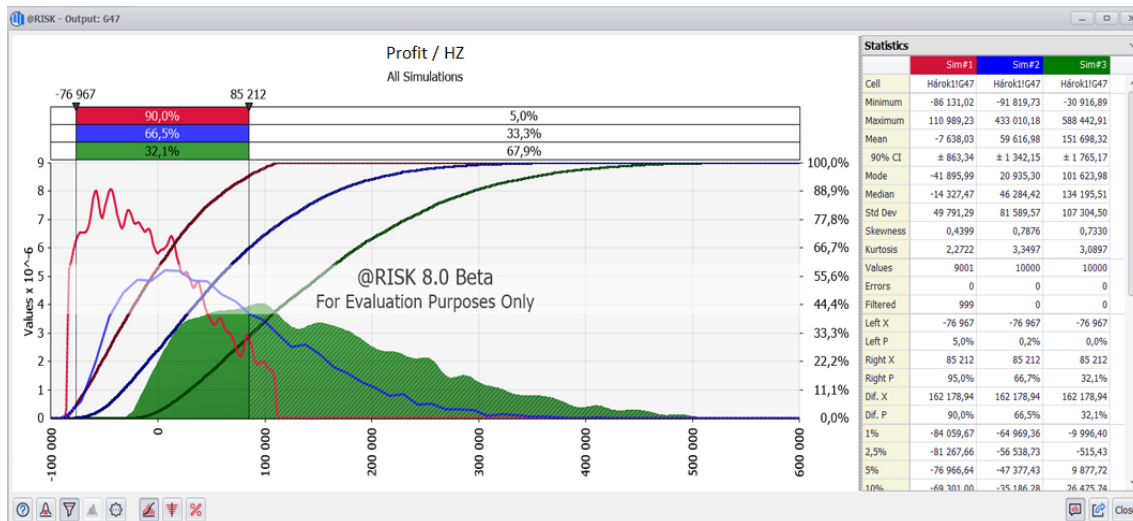
In the case of 40cm and 30cm resolution, a shift to profitability can already be seen in the first year of launching the service. For the purposes of the last simulation, we tested the latest version of the SW Decision Palisade Tools Suite, especially the Beta version of @RISK for comparing functionalities. The table next to the frequencies shows the basic statistics of the resulting simulations of the three scenarios.
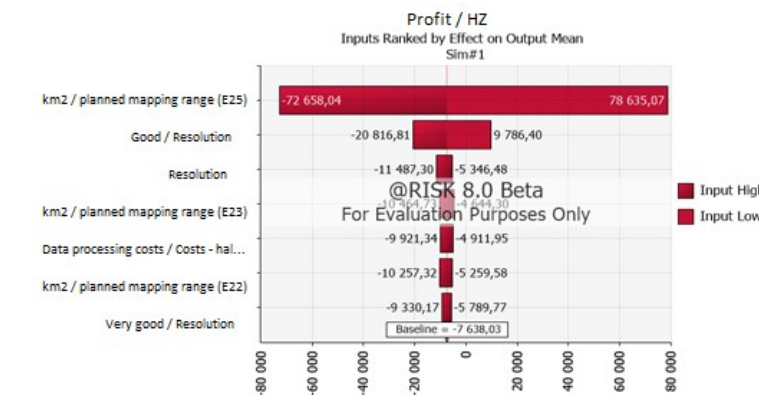

Figure 1 Tornado graph of effect on output

The simulation model also identified the basic most sensitive input parameters, which are described in the table 6 and the tornado figure 1, on the basis of which it is able to set backward scenarios based on adjustments to the desired values at the output.

5. Conclusion

Simulation modelling of business models allows the consideration of risk in data quality, error / accuracy, procurement price for various data parameters and settings of income models and customer portfolios and suitable analysis of potential project development scenarios. To ensure more precise modelling of business value, it is necessary to implement into simulation model crucial variables affecting the precision of the model incorporating different performance approaches (Szabo et al, 2013; Zavadsky and Hiadlovsky, 2014). This model can be continuously modified and refined by extending input raw data from other vectorization and extraction experiments based on learning the convolutional neural network by adding additional data. The understanding of different specific

features of geospatial data which determine through uncertainty, risk and variations different user acceptance is crucial for the development of best practice in related business modelling. Therefore, we are encouraging researchers in business modelling field to enrich the model through identification and calculation of uncertainty related to other geospatial data specifics to provide better business models and data services with higher utility for end users.

6. References

Afuah, A. (2004). Business Models: A Strategic Management Approach. McGraw-Hill/Irwin, New York

Bernknopf, R.L., & Shapiro, C. (2015). Economic Assessment of the Use Value of Geospatial Information. ISPRS Internation Journal of Geo-Information, 4(3), 1142-1165. https://doi.org/10.3390/ijgi4031142

Bernknopf, R.L., Brookshire, D.S., Soller, D.R., McKee, M.J., Sutter, J.C., Matti, J., & Campbell, R.H. (1993). Societal Value of Geologic Maps. United States Government Publishing Office, Washington, DC, USA

Cassadesus-Masanell R., & Ricart J. E. (2011). How to Design a Winning Business Model. Harvard Business Review [online] https://hbr.org/2011/01/how-to-design-a-winning-business-model%20

Castelein, W., Bregt, A., & Pluijmers, Y. (2010) The economic value of the Dutch geo-information sector. International Journal of Spatial Data Infrastructures Research, 5, 58-76. https://doi.org/10.2902/1725-0463.2010.05.art2

Dasgrupta, A. (2013). Economic Value of Geospatial Data: The great enabler. Geospatial World, [online] https://www.geospatialworld.net/article/economic-value-of-geospatial-data-the-great-enabler/

Drucker, P. (1994). The Theory of the Business. Harvard Business Review, [online] https://publicpurpose.com.au/wp-content/uploads/2016/05/Theory-of-the-Business-HBR-Sept1994.pdf

Hunger, D. L., & Wheelen, T. L. (2008). Concepts: Strategic Management and Business Policy. NJ: Prentice Hall.

Krek, A. (2005). Geographic information as an economic good. In book: GIS for sustainable development. CRC Press Taylor & Francis Group, 105-124. https://doi.org/10.1201/9781420037845

Magretta, J. (2002). Why Business Models Matter. Harvard Business Review, [online] hbr.org/2002/05/why-business-models-matter

MIT Libraries. (2019). Geographic Information Systems (GIS): MIT Geodata Repository, [online] cit. 2.3.2020 https://libguides.mit.edu/gis/Geodata

Purchasing & Procurement Center. (2020). Total cost of ownership (TCO) – 3 Key Components of TCO, [online], cit. 8.1.2020 https://www.purchasing-procurement-center.com/total-cost-of-ownership.html

Szabo, S., Dorčák, P., & Ferencz, V. (2013). The significance of global market data for smart e-procurement processes. IDIMT-2013: Information Technology, Human Values, Innovation and Economy: 21th Interdisciplinary Information Management Talks: September 11-13, 2013, Prague, Czech Republic. - Praha : Trauner, 2013, 217-224. ISBN 978-3-99033-083-8

Teece, D. J. (2010). Business Models, Business Strategy, and Innovation. Long Range Planning, 43(2-3), 172-194. https://doi.org/10.1016/j.lrp.2009.07.003

Tolga Saruc N., Dorcak P., & Pollak F. (2013). E-business and its Application in Conditions of Central European Market. Quality Innovation Prosperity, 17(1)

Zavadsky, J., & Hiadlovsky, V. (2014). The Consistency of Performance Management System Based on Attributes of the Performance Indicator: An Empirical Study. Quality Innovation Prosperity, 18(1)

Zott, C., & Amit, R. (2010). Business model design: an activity system perspective. Long Range Planning, 43(2-3), 216-226. https://doi.org/10.1016/j.lrp.2009.07.004

POSITION OF PUBLIC PROCUREMENT IN SUPPLY CHAIN

Paulína Kaselyová, Slavomíra Stašková, Marek Meheš, Jozef Lukáč

Department of Corporate Financial Management

The Faculty of Business Economics with seat in Košice,

The University of Economics in Bratislava

paulina.kaselyova@student.euke.sk, slavomira.staskova@euke.sk

marek.mehes@euke.sk, jozef.lukac@euke.sk

Keywords

public procurement, supply chain, design of experiment, cost

Abstract

This contribution provides methodological instructions for business entities how to increase the chances of success when applying for public contracts in the public procurement. To create a model, we performed an ex-post analysis of existing data on public procurement to identify general factors that influence the process of public procurement. Subsequently, we used the method of design of experiment and examined, which factors are significant in the process of public procurement for selected type of good.

1. Introduction

Public procurement is practically the only standardized mean in the supply chain for the acquisition of goods and services that public sector cannot provide itself. Despite this fact, not enough attention is currently paid to measuring the economic efficiency of public procurement processes. In this context, savings are most often mentioned, which are quantified as the difference between the estimated value of the contract and the final price, while investigating the factors with significant impact on these savings. This is also mentioned by the author Pavel (2010), who states that the efficiency of public procurement can be assessed according to the final price of the public contracting authority. The aim of public sector organizations in public procurement should be to achieve the lowest possible price with the intention of the lowest possible burden on public budgets.

The price itself is currently the centre of attention in measuring the effectiveness of public procurement, and to the price is given the greatest attention in studies dealing with this issue. An example may by a study by Gregor and Nemec (2015), who examined savings in public procurement as the difference between estimated and final price. The authors state that in their research sample they identified 3 statistically significant factors, that have an impact on savings in public procurement – the number of bidders, the use of EU structural funds and the use of the lowest price criterion. The authors state that each additional bidder in public procurement increased the overall savings by an average of 2.63 %. In the case of financing the contract from EU resources, on the other hand, the savings decrease by an average of 1.54 %, and if the only criterion in a public procurement is the

lowest price, the savings of such procurement increase by an average of 1.06 %. Several similar studies have been performed, such as by Soudek, Skuhrovec (2013), Šipoš and Klatik (2013), Ilke, Rasim and Bedri (2012), Gómez-Loba and Szymanský (2001), by Pavel (2010), Bajari (2002), Nemec, Vítek, and Meričková (2005) and others, who evaluate the effectiveness of public procurement based on price and the factors that determine it. However, this is insufficient, because in the process of public procurement there are several factors that determine it and that affect its final economic efficiency. Also, most of the attention in examining the effectiveness of public procurement is paid only to the contracting authority – institutions of the public sector. There is no interest in measuring the performance of companies applying for public contracts in public procurement. For these reasons, in the article we will focus on examining the effectiveness of public procurement processes from the perspective of companies applying for public contracts, to determine the factors, which increase the probability of success in public procurement of specified good. The result of our research is a model that should increase the quality of decision-making of the business entity when applying for public contract.

2. Methodology of the research

To obtain primary data, we focused on the process of public procurement in Slovakia in electronic contracting system. Within the electronic contracting system, we collected data on public contracts of one selected product – PC desktop. The basic data set represented 504 desktop procurement contracts. To examine the standardized product, we have defined in advance the technical parameters of contracted computers. After reviewing the above number of contracts, we created a research sample of data that included 170 public contacts of PC desktops.

For processing collected data we decided to use the method of design of experiment (DoE). In the method of design of experiment, changes in the process are implemented and appropriate responses are observed to obtain information about the given process, which help to improve this process. According to the Montgomery (2001), the design of experiment represents a test, respectively sequence of tests in which the input factors of the process are purposefully changed to examine the corresponding changes of the output variable, the so-called response. Experiments are considered a very important tool in quality management and represent one of the main activities in improving any process, for example about the application of Six Sigma principles.

In general, we can state that many variables affect each process that we intend to optimize. Variables, that we can influence are referred to as factors and can be qualitative or quantitative in nature. The lower (-1) and upper (1) levels are determined for each factor. At the end of each process there are some outputs and such a measurable output is called a response. It is a variable, that is used to express the results of the experiment and at the same time represents a factor-dependent variable. The main task of the DoE method is to find such a combination of factors, that have the greatest influence on the process and to design levels of these factors so that the response value is as favorable as possible and the process is fully functional and efficient (Miller, 2010).

When examining processes with many factors, it is often not possible to carry out a full factorial design of experiment due to lack of time, resources or human capacity. If we assume that higher-order interactions are not relevant, then knowledge of the main effects and second-order interactions can be obtained through the partial factorial experiment, that is most common used in practice. This type of experiment must be abbreviated in such a way that all significant interactions of the chosen model can be quantified. This plan of experiment has the form 2^{k-p}, where k represents the number of factors and $\frac{1}{2}^p$ represents part of the complete factorial experiment 2^k. The number of partial experiment trials n must not be less than the number of factors, it means $n \geq k$. Partial factorial experiments can thus be divided e.g. to half partial 2^{k-1}, quarter partial 2^{k-2}, octave partial 2^{k-4}, etc.

The quantification of effects and mutual interactions and the construction of a regression model are performed like the complete multifactorial experiment. Advantages of partial factorial experiment are mainly examination a relatively large number of factors with the implementation of a smaller number of measurements, or monitoring the factors within wide range of conditions. On the contrary, limitations are the susceptibility to special cases of variability and to missing data.

3. The results of the research

To examine the impact and interactions of the variables that appear in the procurement process for the selected type of good, we decided to use the method of design of experiment. The main goal of the experiment was to determine, which variables have the most significant influence on the public procurement process. Based on a model case, our intention was to provide businesses with instructions on how to proceed in the public procurement process to increase their effectiveness. For the needs of our research, we decided to use a two-level factorial experiment, with which we can examine several factors at once. When planning the experiment, we chose 8 factors to examine their effects on the selected response. The number of measurements of the complete factorial design of experiment was based on the relation 2^k, where k is the number of factors, which in our case would mean performing up to 256 measurements. For this reason, as well as due to lack of time, capacity and resources, we decided to apply a partial factorial (eighth: 2^{8-3}) design of the experiment for 32 measurements. However, even the reduced number of measurements allowed us to examine some interactions between the selected variables.

In the selection process, we have chosen such input variables for which the effect on the output variable can be assumed, it means response which can be measured. We specified the factors investigated in the experiment based on ex-post analysis of realized public procurement of desktop computers. It is advisable to consider the experiment with as large as possible number of variables, even though the overall complexity of the experiment increases. Initially, we considered 12 factors, which proved to be unfeasible, as their combinations in the experiment design did not fit our data set. For this reason, by gradually eliminating the least causally significant factors, we have created set of eight factors, based on which we compiled the experimental plan. For the purposes of our paper, we decided to examine the factors we have chosen at two specific levels - lower and upper. Specifically, we focused on the factors described in Table 1, including their lower (-1) and upper (1) levels.

Table 1 Characteristics of the investigated factors; Source: own research

Factor	Description of the factor	Lower level (-1)	Upper level (1)
Institution (A)	Type of contracting authority by source of funding	MR (municipal budget)	ŠR (state budget)
Location (B)	The same territorial region of the contracting authority and the applicant	IK (different region)	TIK (same region)
Number of contracted desktops (C)	Number of desktops	3 (1 – 4)	5 (5 and more)
Auction (D)	Possibility of electronic auction	A (Yes)	N (No)
P1 (E)	Technical parameter – processor type	i3	i5
P2 (F)	Technical parameter – RAM size	4 GB	8 GB
Delivery of goods (G)	Number of days needed for delivery of goods	1 (1 – 2 days)	3 (3 and more days)
Estimated value (H)	Maximum amount of public financial resources	400 (up to 499 Eur)	600 (500 – 600 Eur)

After selecting the type of experiment, setting the basic principles, determining the individual factors including their lower and upper levels, we could prepare an experiment plan. Subsequently, using the MiniTab program, we generated random combinations of values of selected factors within 32 measurements. Based on the factor plan, we filtered public contracts from the set of realized public procurement using electronic contracting system according to the generated combinations of values of selected factors. The result of this selection was a set of 32 successful candidates whose contracts met combinations of factor values from the experimental design. Subsequently, we simulated the conditions of electronic contracting system to obtain feedback from companies that have used the public procurement process in the past to sell their goods (desktop computers).

3.1. Analysis of experiment results for desktop computers

To separate significant factors from insignificant ones, we used a Pareto diagram and set the level of significance at the level $\alpha = 0.05$. From graph 1, the most significant effect on the response (price) has the type of institution. Other factors which *p-value* is below the selected significance level are location and technical parameter P1. Combinations of factors type of institution and parameter P2 and type of institution and parameter P1 are also significant.
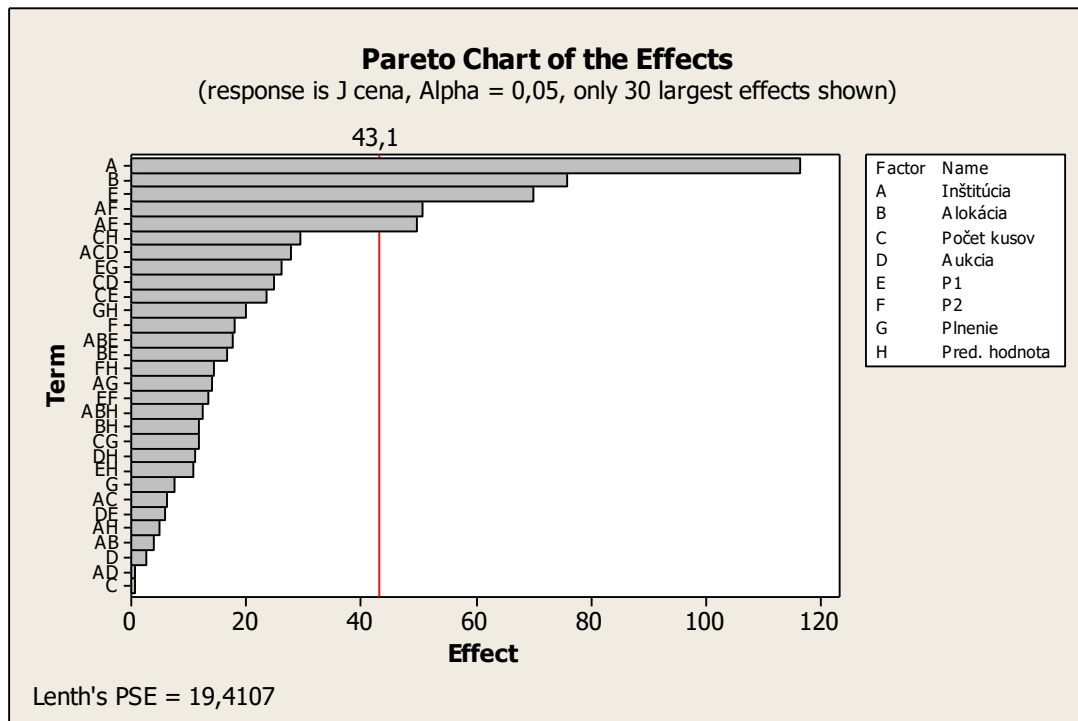

Chart 1 Pareto diagram; Source: own research

Based on the probability graph of the normal distribution (chart 2), the values approaching the regression line represent insignificant factors and are probably noises. On the other hand, the red dots represent significant factors influencing the investigated response.
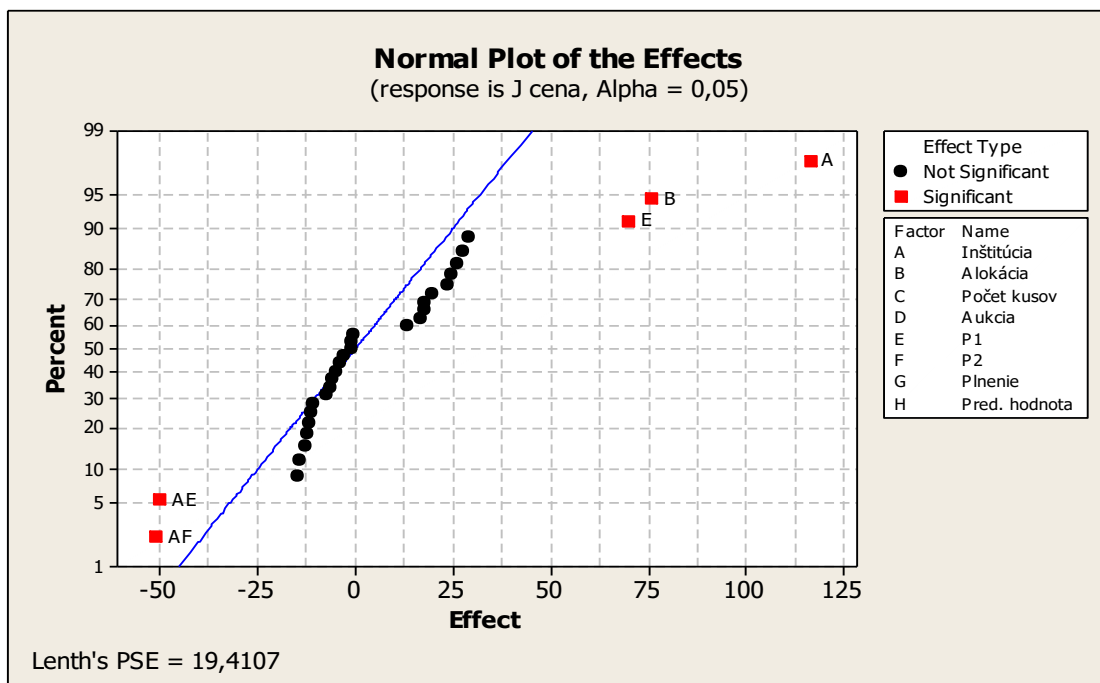
Chart 21 Probability graph of the normal distribution; Source: own research

In the next section, we examine the main effects of factors for desktop computers. It can be seen from Graph 3, that we rank the institution (A), location (B) and parameters P1 (E) and P2 (F) among the factors with a positive effect, because the lines of these factors rise from left to right.
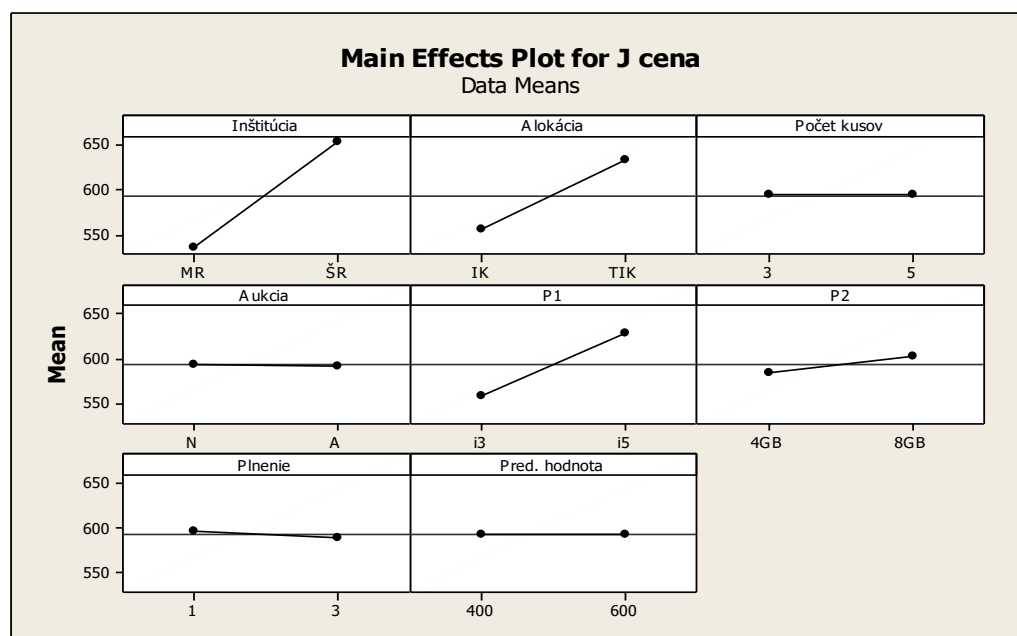

Chart Legend	
Inštitúcia – Institution (A)	*Parameter P1 (E)*
Alokácia – Location (B)	*Parameter P2 (F)*
Počet kusov – Number of contracted desktops (C)	*Plnenie – Delivery of goods (G)*
Aukcia – Auction (D)	*Pred. hodnota – Estimated value (H)*

Chart 3 Main effects; Source: own research

Even in this case, we consider only three of main effects to be statistically significant, namely the type of institution (A), the location (B) and the technical parameter P1 (E). On the other hand, the factors – number of contracted goods (C), auction (D), delivery of goods (G) and estimated value (H) have no effect on the response (price).
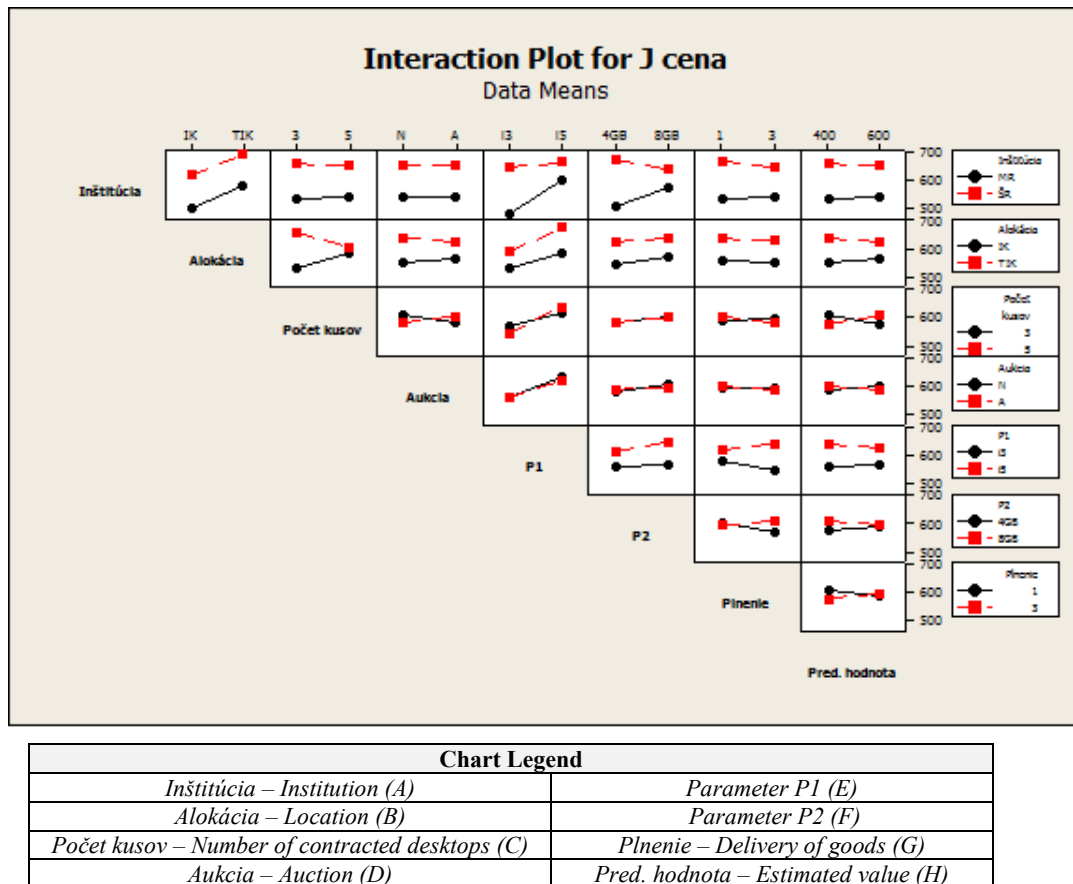

Chart Legend	
Inštitúcia – Institution (A)	*Parameter P1 (E)*
Alokácia – Location (B)	*Parameter P2 (F)*
Počet kusov – Number of contracted desktops (C)	*Plnenie – Delivery of goods (G)*
Aukcia – Auction (D)	*Pred. hodnota – Estimated value (H)*

Chart 4 Mutual interactions; Source: own research

We also examined the effect of mutual interactions of factors on the response. Based on Graph 4, we conclude that strong and moderate interactions appear between the following factors:

- institution (A) and parameter P1 (E), institution (A) and parameter P2 (F),

- location (B) and number of contracted goods (C),

- number of contracted goods (C) and auction (D), number of contracted goods (C) and parameter P1 (E), number of contracted goods (C) and estimated value (H),

- parameter P1 (E) and delivery of goods (G),

- parameter P2 (F) and delivery of goods (G), parameter P2 (F) and estimated value (H),

- delivery of goods (G) and estimated value (H).

4. Conclusion

After statistical data processing through the MiniTab program, we concluded that to maximize the effectiveness of companies applying for public contracts should focus on those factors that increase

the chances of success in the public procurement. For the observed group of goods, we state that a total of five factors have a statistically significant and positive effect on the price – namely the type of institution (A), location (B), parameter P1 (E), combination of factors type institution (A) and parameter P1 (E) and combination of type of institution (A) and parameter P2 (F). We further state that to maximize the price of the public contract of desktop computers, companies should focus on public contracts initiated by a state-funded institution from the same region as the successful tenderer and that require the i5 processor type. We also assumed positive influence of the factors number of contracted goods (C), auction (D), delivery of goods (G) and estimated value (H) on the price, but the results showed that these factors have a minimal impact and can be considered statistically insignificant. The main effect was also positive for factor P2 (F), which is not statistically significant. The statistical significance of the mutual interactions was confirmed only for the combination of institution (state budget) and parameter P1 (i5) and the combination of institution (state budget) and parameter P2 (4GB).

5. References

BAJARI, P. (2001). Comparing Competition and Collusion in Procurement Auctions: A Numerical Approach. Economic Theory, 18(1), 187-205.

GÓMEZ-LOBO, A., & SZYMANSKI, S. A (2001). Law of Large Numbers: Bidding and Compulsory Competitive Tendering for Refuse Collection Contracts. The Review of Industrial Organization, 18(1), 105-113.

GREGA, M., & NEMEC, J. (2015). Factors Influencing Final Price of Public Procurement: Evidence from Slovakia. Procedia Economics and Finance 25 (1), 543-551.

ILKE, O., & RASIM, O., & BEDRI K. (2012). Public Procurement Auctions and Competition in Turkey. Review of Industrial Organization, 40(1), 207 – 224.

MILLER, I. (2010). Návrh a analýza experimentu s pomocí MINITAB, Praha, Interquality, 2010, ISBN 978-80-902770-5-2.

MONTGOMERY, D. C. (2001). Design and Analysis of Experiments. 5th edition. North Carolina, USA: SAS Institute Inc., 2001, ISBN 0-471-31649-0.

NEMEC, J., & VÍTEK. L., & MERIČKOVÁ, B. (2005) Contracting-out at local government level: theory and selected evidence from the Czech and Slovak Republics. Public Management Review, 7(4), 637-648.

PAVEL, J. (2010). Analýza vlivu míry konkurence na cenu rozsáhlých staveb dopravní infrastruktury. Politická ekonomie, 58(3), 343-356.

SOUDEK, J., & SKUHROVEC, J. (2013). Public procurement of homogeneous goods: The Czech Republic case study. IES Working Paper 05/2013. IES FSV. Charles University, Online http://ies.fsv.cuni.cz/sci/publication/show/id/4833/lang/en

ŠÍPOŠ, G., & KLÁTIK, P. (2013). Kvalita verejného obstarávania na Slovensku. Transparency International, Online: http://www.transparency.sk/wp-content/uploads/2013/01/TIS-Analyza-VO-2012-1.pdf

CORPORATE REPUTATION OF FAMILY OWNED BUSINESSES- PARENT COMPANIES VS. THEIR BRANDS

František Pollák, Peter Dorčák, Peter Markovič

Faculty of Business Management
University of Economics in Bratislava
frantisek.pollak@euba.sk

Keywords

Reputation management, Brand development, Information distribution, Family business, Business succession.

Abstract

The main aim of the paper is to discusses the issue of innovative approaches to the reputation management of the top family owned businesses and their brands. As the family business issue has many applications in managerial and economic sciences, we have decided to examine one of its parts, in the form of a perception on the issue of the reputation of the biggest players on the global market. From previous researches, we have found that in selected cases there is a significant difference between the distribution of information or in general in marketing communication of the parent company and the marketing communication of the brand owned by the company. This creates the place for a more in-depth analysis, which we then dealt with in the present study. The findings of the analyzes will provide a better insight into the issue of sustainable brand development, by identify good practices, as well as highlighting weaknesses in the form of unclear communication of some parent companies in our sample.

1. Introduction

In terms of reputation, the literature offers a wide range of views, from interpretational formalized views to views of an almost informal nature. In general, all of these views agree that reputation as a business asset is an extremely fragile element. At the same time, we are confronted with the claim that reputation is an element that every organization has to offer without distinction (Marsden, 2013). Unlike trust, often described as an oriented state (Delina, Drab 2010), reputation is a more complex term, but we also encounter the claim, in our opinion not quite accurate, that it can be seen as a synonym for trust and reputation. Let's look at how reputation is defined within the scope of the available literature. Reputation is perceived as the overall quality, the optics, how the organization is perceived or judged by individuals (Griffin, 2008). The definition is a fairly simple interpretation; on the other hand, the purely formal encyclopaedical definition (Brittanica, 2017) deals with reputation, again as a general quality, this time however it extends this quality to the very nature of the organization, which is clear and known to the target audience, with emphasis being placed on how these audiences perceive the attitudes, actions and opinions of the organization. From the point of view of corporate management, it can be argued that, in the past, reputation was the domain of marketing, while currently it is integrated into the company strategy itself (Cooper et al., 2012). An

interesting management view of the issue of reputation is the claim (Benedigová, 2003) that reputation can also affect activities directly unrelated to the fulfilment of corporate goals. It may be informal expressions of personnel, insufficient or unsatisfactory responses to customer complaints, or ill-considered statements of representatives of organizations for the media. From the point of view of history (Burke, Graeme, Cooper, 2012), it is possible to believe that reputation or reputation management evolved from public relations. As mentioned above, reputation is considered by many authors to be an asset of an organization. We also encounter the claim that it is the most comprehensive business asset (Helm et al., 2011). Despite the challenge in the form of organized corporate reputation management and the use of comprehensive measurement approaches, reputation is still a relatively unexplored area (Delina 2014, Dorčák, Markovič, Pollák 2017, Soviar et. Al 2019, Markovič, Dorčák, Pollák 2019).

2. **Objectives and methods**

The primary aim of our research is to discusses the issue of innovative approaches to the reputation management of the top family owned businesses and their brands. The issue of managing online reputation as a new phenomenon in the form of fragile intangible assets is gradually gaining on importance and it is becoming one of the essential prerequisites for responsible and sustainable reputation management. From previous researches, we have found that in selected cases there is a significant difference between the distribution of information or in general in marketing communication of the parent company and the marketing communication of the brand owned by the company (Pollák et al. 2019). This creates the place for a more in-depth analysis, which we then dealt with in the present study. For the purpose of in-depth analyses, we choose significant sample from chosen group of subjects, namely the World's Top 10 family owned businesses (Bain In: Family Capital, 2019), subsequently, we choose one top brand from the portfolio of each subject in our sample. Within the testing, we considered the entire spectrum of perceiving reputation of our subjects, since we compared the whole specter of relevant virtual factors and connections measured by us against significant and relevant ranking of the mortar world provided by Family Capital; the online publishing company dedicated to the global family enterprise sector, based on y. 2017 revenues in USD. We expected, that the findings of the analyses will provide a better insight into the issue of sustainable brand development, by identify good practices, as well as highlighting weaknesses in the form of unclear communication of some parent companies in our sample. Within our research, we used the methodology of multi-factor analysis of online reputation, namely its modified version TOR, which we describe in more detail in our previous studies (Dorčák, Markovič, Pollák, 2017, Markovič, Dorčák, Pollák 2019), and which we would like to present as a priority possibility to measure the entire power of online reputation of a subject in the Internet environment. The measurement itself takes place in three steps, in first step, based on the chosen methodology, it is necessary to analyze the sentiment of the first ten keyword search results, namely own or established name of the particular subject through Google search. To increase the validity of retrieved data, the sentiment of the first ten search results of own name of specific subject is parallelly tested by its own name in Google tab News. This parameter is displayed in our study as ASA (Dorčák, Markovič, Pollák, 2017). In the second step, we identify the determinants of online reputation, for the purposes of our study, we will call them reputators. As a reputator, we can identify any determinant which has the ability to influence the perceived online reputation of a subject, at the same time, it can be quantified, and it is possible to express its value in percentages. Given the various business fields entities operate in, reputation determinants cannot be clearly defined in advance. In terms of advantage quantification, it is possible to approach individual reputation determinants by calculation of reputators' competitive score – the number of users/fans/customers/followers of the particular entity (Socialbakers, 2019) has relative to the sum of all tested subjects. The result serves as a basis for calculating the percentage of the

reputators' competitive score of the particular entity, this parameter is displayed in our study as CS (Dorčák, Markovič, Pollák, 2017). In the third step, we can proceed to the actual calculation of the overall power of online reputation of a specific TOR subject which then serves as a starting point for complex comparison of overall power of reputation across all analyzed subjects. In this case, the value of the overall online reputation is an arithmetic average of individual indicators (partial score according to individual determinants). Relationships between factors were examined and results were visualized in order to provide better infight to the solved topic.

3. Results and discussion

Using the advanced sentiment analysis, we calculated partial score for each of selected brands from the portfolio of entities based on the nature of the first 10 Google search results. Google and its search results are, however, only one of many ways in which potential customers can access relevant information. Considering the previous research in the field, we identified the following other determinants of online reputation (reputators), in particular:

- Facebook (total number of followers of the official and verified global profile);

- Twitter (total number of followers of the official and verified global profile);

- YouTube (total number of subscribers of the official global profile);

- LinkedIn (total number of followers of the official global profile).

Each of these reputators has its own system which determines the overall score. But basically, all of them operate with a certain tribe of the customers (followers, fans, subscribers). For the purposes of further analysis scores of partial reputators were unified to the parameter which we named competitive strength, hereafter referred to as CS and converted into a percentage. Before we present our own interpretations, we consider necessary to present partial evaluations of the subjects as well as all relevant indicators in one summary table:

No. *	Subject/ Result sentiment **	ASA Score (%)	FC Score (%)	TC Score (%)	YC Score (%)	LC Score (%)	Number of pages indexed by Google	TOR Score (%)
1.	Sam's Choice (Walmart Inc.)	28	9	3	1	6	18,800,000	9
2.	Audi (Volkswagen AG)	70	2	52	7	34	1,400,000,000	33
3.	Berkshire Hathaway Inc.** *	21	0	0	0	2	41,900,000	5
4.	Fiat (Exor NV)	-33	10	4	8	27	731,000,000	4
5.	Lincoln (Ford Motor Company)	30	3	3	29	0	665,000,000	13
6.	Lidl (Schwarz Gruppe)	46	0	0	1	1	108,000,000	10
7.	Mini (BMW AG)	36	31	7	18	2	4,150,000,000	19
8.	Wilbur Chocolate (Cargill, Inc.)	67	0	0	0	0	2,810,000	13
9.	Jaguar (Tata Sons Ltd)	55	45	31	36	22	610,000,000	38
10.	Georgia-Pacific (Koch Industries, Inc.)	36	0	0	1	5	423,000,000	8

Table 1: Overall (Total) online reputation, Source: Own processing

Legend: * Sorted by order according to Family Capital ranking, based on y. 2017 revenues in USD.

**For better interpretation in subsequent visualizations, the results were rounded to the nearest whole number, therefore, the value of 0 does not necessarily indicate the absence of the subject's profile on

the social network, it may also indicate a very low representation from the point of view of the size of the tribe of followers with respect to competing entities.

 ***As it is a specific subject, we left it in its original form for analysis purposes.

In the first step, we will look at the results of an extended analysis of the sentiment of ten selected brands belonging to the portfolio of the largest family businesses in the world. The first brand in the series is Sam's Choice, a brand from the portfolio of retail giant Walmart Inc. From the point of view of the brand's virtual presence in the first 10 google search results, the brand's reputation is relatively vague compared to that of the parent organization; From the point of view of the second observation, the situation was even more embarrassing, all the references in the news category were irrelevant, which in the end meant ten neutral sentiments, with the gain only a little more than ten percent of the total potential. Here, the betrayal of neutral sentiment is fully demonstrated, which, although it does not harm the brand, but can cause significant differences in rankings when compared within a larger sample of active subjects. The second analyzed subject was Audi, a brand from Volkswagen's portfolio. Although it did not surpass its parent organization in the general search results category, anyhow it gains the full number of points in the search results in the news category, as part of our observation, it is a demonstration of exceptional media work of the brand managers, which also helped the brand to be number one among entities in terms of reputation based on search results. Third in line was Warren Buffett's company Berkshire Hathaway, this company was specific in our measurement, as it does not directly have its own brands, we left it as a reference in our table of results. The fourth in a row was the Fiat brand belonging to the portfolio of Exor NV. Here we will stay, that in comparison with the competing brand Audi, in the first observation the Fiat brand achieved only slightly worse numbers, but in the second observation in the category of news, it was the exact opposite of the Audi brand, where it recorded negative sentiment in each of the ten places in search. This earned Fiat the last place among the analyzed subjects within the given category. The fifth in a row was the Lincoln brand, which belongs to the Ford Group's portfolio. Except for small differences, in the first observation it recorded the same values as its competitors, in the second observation it completely copied the Sam's Choice brand from Walmart's portfolio, where it recorded neutral sentiment in each position. At this point, a trend began to emerge, temporally called the black hole effect, when a strong parent company attracts all the attention with its gravity and makes it significantly impossible for the subsidiaries in the portfolio to grow on their own. Sixth is Lidl, which belongs to the Schwarz Gruppe portfolio. In the first observation, the brand reaches only a little more than 30% of the potential, in the second observation, the excellent result thwarted the occurrence of negative sentiment in the third position in the search results. Seventh in the ranking is the brand MINI belonging to the BMW portfolio. From all sample, MINI is one of two brands, that in the first of the observations scored with more than 90 points, which is equivalent to 60% of the total possible reputation. The chances of a good position were thwarted by the presence of negative sentiment in the third position in the news category, which dealt with the closure of the local dealership, as well as the amount of irrelevant content probably caused by the general nature of the brand name. The eighth in the ranking was the Wilbur chocolate brand, which belongs to the Cargill portfolio. With a gain of more than 60% of the total potential, Wilbur ranked first in this round of observations. Also, in terms of the overall result within the ASA parameter, Wilbur significantly exceeded the reputation of its parent company Cargill. Here we began to record another of the trends, temporally called the celebrity effect, when a huge corporation focusing out of the mainstream complex (and thus) uninteresting media coverage is overshadowed by a relatively small brand with a solid tradition and a strong mainstream product. The ninth in a row was the Jaguar brand, which belongs to the portfolio of the Indian company Tata. It was the first subject to record almost identical results in both observations. Once again, we are seeing the aforementioned celebrity effect, where the Jaguar brand, the winner of the overall standings, significantly exceeded the level of reputation of its parent company Tata. Our observations are closed by the Georgia-Pacific brand, which belongs to the Koch

Industries portfolio. Based on the observations, it can be stated that the results of a predominantly organic nature did not show signs of significant reputational error. In any case, the organic nature of the results left the subject to chance, which was reflected in the negative sentiments recorded in the first, second and fourth place in the search category of the news category. This was also due to the fact that the brand only slightly exceeds the potential of the parent company. All relevant data is graphically processed in subsequent Figure number 1, part a).

(a)   (b)
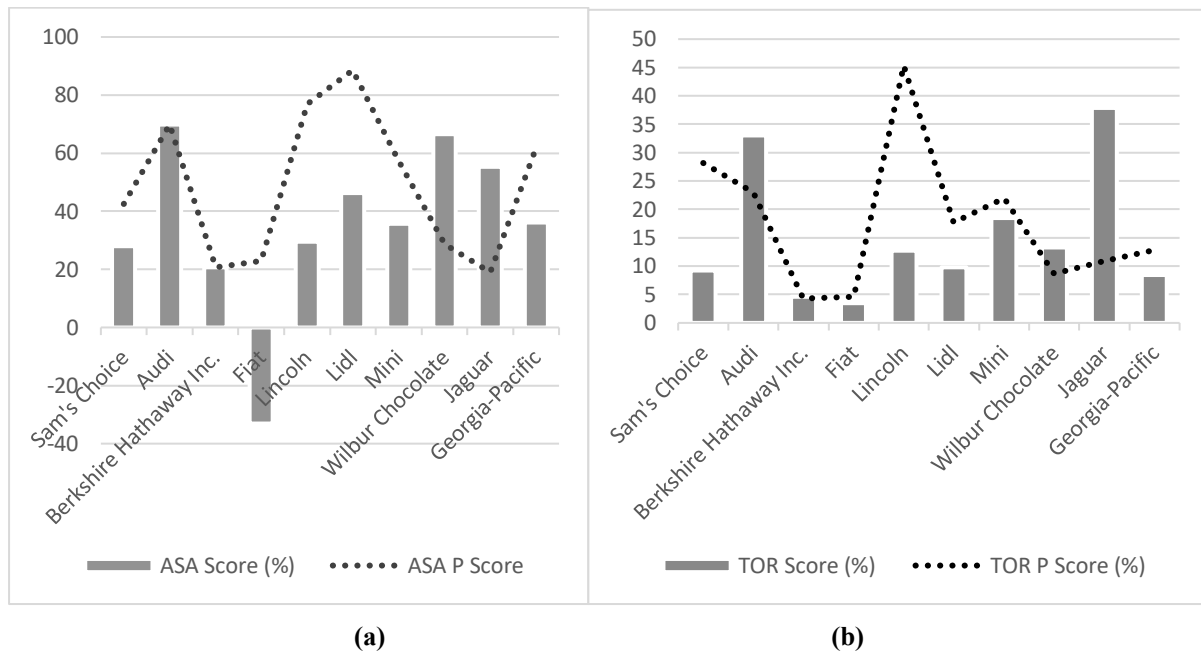
Figure 1: Reputation of brands vs. reputation of parent companies, Source: Own processing

From the point of view of comparing the size of the tribes of individual brands in the environment of virtual social networks and virtual social media, the situation was diverse. The brand with the largest tribe on the most popular virtual social network Facebook is Jaguar. In terms of the number of fans, it has almost half size of the total tribe compared to the sum of the sizes of the tribes of all other brands in our analysis. This finding is not surprising, in any case, it points to the celebrity effect already mentioned in our study, especially considering the fact, that parent company Tata jumped from the numbers on the level of statistical deviation through its Jaguar brand into the first place in our test sample. When it comes to the virtual social network / virtual social media Twitter, the market is dominated by the Audi brand followed by the Jaguar. YouTube once again shows the trend of a black hole effect, where the Ford dominated in the parent company environment with almost 75% of the market, the subsidiary Lincoln with a more modest 29% leaves the first place in the brand environment to the Jaguar. The professional social network LinkedIn is divided almost equally between three brands Audi, Fiat and Jaguar.

From the point of view of the overall results visualized in part b) of Figure 1, it can be seen that the overall level of reputation of their parent companies is surpassed by only three entities, namely- Audi, Wilbur Chocolate and Jaguar. Jaguar, with a gain of almost 40% of the total possible reputation level, also becomes the winner in our comparison. In terms of the overall level of reputation of its parent company, it surpasses it more than threefold.
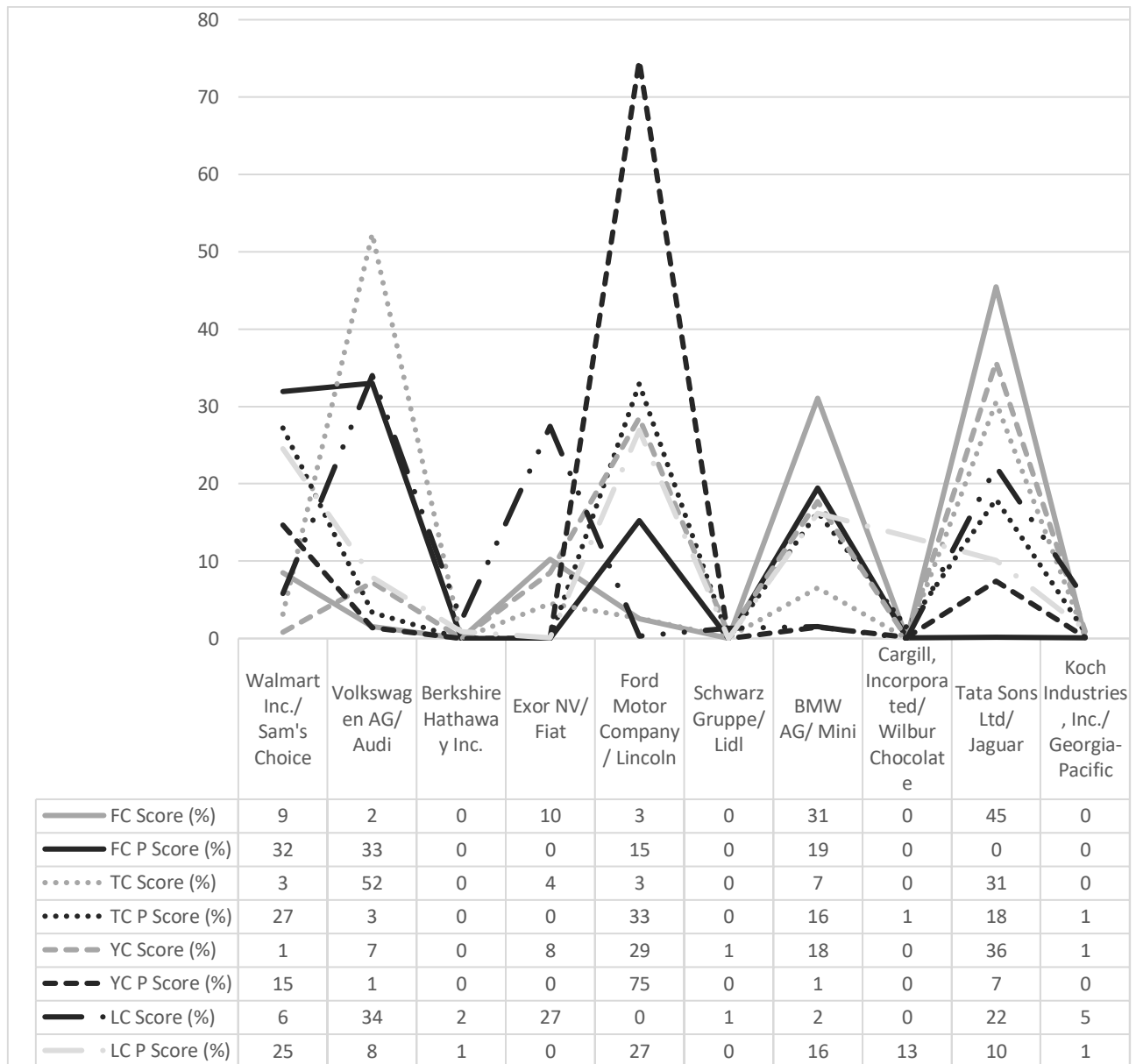
	Walmart Inc./ Sam's Choice	Volkswagen AG/ Audi	Berkshire Hathaway Inc.	Exor NV/ Fiat	Ford Motor Company / Lincoln	Schwarz Gruppe/ Lidl	BMW AG/ Mini	Cargill, Incorporated/ Wilbur Chocolate	Tata Sons Ltd/ Jaguar	Koch Industries, Inc./ Georgia-Pacific
FC Score (%)	9	2	0	10	3	0	31	0	45	0
FC P Score (%)	32	33	0	0	15	0	19	0	0	0
TC Score (%)	3	52	0	4	3	0	7	0	31	0
TC P Score (%)	27	3	0	0	33	0	16	1	18	1
YC Score (%)	1	7	0	8	29	1	18	0	36	1
YC P Score (%)	15	1	0	0	75	0	1	0	7	0
LC Score (%)	6	34	2	27	0	1	2	0	22	5
LC P Score (%)	25	8	1	0	27	0	16	13	10	1

Figure 2: Social media strength of brands vs. their parent companies, Source: Own processing

Figure number 2 shows the sizes of the tribes of individual brands and their parent organizations in the environment of dominant virtual social networks and virtual social media. Two possible trends were identified, the first was called the black hole effect (the case of Walmart and Ford), the second was called the celebrity effect (the case of Wilbur and Jaguar). In the last figure, the index "P" indicates the value of the parent organization of which brand, the values of the indicators of the parent organization are marked in a dark shade for better visibility. As we mentioned in the legend for Table 1, the value of 0 does not necessarily indicate the absence of the subject's profile on the social network, it may also indicate a very low representation from the point of view of the size of the tribe of followers with respect to competing entities.

4. Conclusion

Based on the findings, we can conclude as well as in our previous studies that sustainable development of corporate reputation management combines offline and online techniques, as the both

worlds are connected. Even though different environments require specific approaches, recorded different dynamics, and required specific tools. It is almost impossible to be a star only in one world. However, by its nature, online reputation is more fragile. Eliminating negative publicity while maximizing positive media outputs in opinion-forming Internet media, and the displacement of neutral or negative search results to irrelevant positions presented by the second to n-th side of Google search results is the key approach to online reputation management. From the point of view of fulfillment of the goal of our study, namely investigate the difference between the distribution of information of the parent company and the marketing communication of the brand owned by the company, we realized in-depth analysis, with the findings proposes existence of three scenarios. Scenario of shared reputation, when the brand and the parent company sharing more less same reputation. Scenario of the black hole, when the mother company creates a media event horizon that absorbs all media light. And finally, the third scenario, when a big mother corporation focusing out of the mainstream complex, and thus uninteresting media coverage is overshadowed by its relatively small brand with a solid tradition and a strong mainstream product. This scenario also leaves the parent company most vulnerable in terms of reputation development. As it creates conditions for asymmetric distribution of information. The findings of the analyzes will provide a better insight into the issue of sustainable brand development, by identify good practices, as well as highlighting weaknesses in the form of unclear communication of some parent companies in our sample. From the point of view of research limitations, it is possible to mention the limited sample size as well as the qualitative data processing. Given the nature of the analysis, the entity consisting of a highly relevant sample and the very purpose of the analysis, which is in the nature of discovering specificities and creating preconditions for further quantitative research, we consider the limitations of our research to be acceptable.

5. Acknowledgement

6. References

Bain, D. (2019). The World's Top 750 Family Businesses Ranking. Available online: https://www.famcap.com/the-worlds-750-biggest-family-businesses/ (accessed on 05 June 2019).

Benedigová, M. (2003). Reputáciu firmy nemožno ponechať len na PR manažéra. Available online: http://www.etrend.sk/podnikanie/reputaciu-firmy-nemozno-ponechat-len-na-pr-manazera.html (accessed on 27 January 2016).

Burke, J. R., Graeme M., Cooper, L. C. (2012). Corporate Reputation: Managing Opportunities and Threats. London: Gower Publishing, 2012. p. 356

Brittanica. (2017). Encyclopedic Dictionary. Available online: https://www.britannica.com (accessed on 01 June 2019).

Cooper, C. et al., (2012). Corporate Reputation: Managing Opportunities and Threats. London: Gower Publishing. ISBN 978-14-09460-39-8

Delina, R. (2014). Transparency in Electronic Business Negotiations – Evidence Based Analysis. Quality Innovation Prosperity. 2014. Volume 8. pp. 79 – 89.

Delina, R., Drab, R. (2010). Socio-economic Aspects of Trust Building for the Elec-tronic Business Platforms. E & M Ekonomie a management, 2010. Volume 13. pp 110 – 122.

Dorčák, P., Markovič, P., Pollák, F. (2017). Multifactor analysis of online reputation as a tool for enhancing competitiveness of subjects from automotive industry. In: Ekonomický časopis. Vol 65, No. 2 (2017), pp. 173-186. ISSN 0013-3035

Griffin, A. (2008). New Strategies for Reputation Management: Gaining Control of Issues, Crises & Corporate Social Responsibility. Kogan Page Publishers: London. UK. 2008.

Helm, S., et al. Reputation Management. Springer-Verlag: Berlin, Germany, 2011.

Markovič, P., Dorčák, P., Pollák, F. (2019). Online Reputation Management. 1st Edition. Praha: Professional Publishing, 2019. 204 p. ISBN 978-80-88260-35-6.

Marsden, H. (2013). Guard Your Reputation On-line. Smartebookshop: Birmingham. UK. 2013.

Pollák, F. et al. (2019). Selected Approaches to Sustainable Development of Corporate Reputation - The Reputation of Family Owned Businesses in the Face of Turbulent Environment. Global Business Conference 2019 Proceedings: Doing Business in a Turbulent Environment. Zagreb, Croatia: Institut za inovacije, 2019, pp.128-138. ISSN 1848-2252.

Socialbakers. (2019). 2019 Social Media Marketing Statistics. Available online: http://socialbakers.com. (accessed on 1 June 2019).

Soviar, J., et al. (2019). The Presentation of Automotive Brands in the On-Line Environment—The Perspective of KIA, Peugeot, Toyota and VW in the Slovak Republic. In: Sustainability 2019, 11(7), 2132; doi:10.3390/su11072132

CRISIS MANAGEMENT OF CORPORATE REPUTATION- ANALYSIS OF SELECTED E-COMMERCE ENTITIES IN TIMES OF GLOBAL PANDEMICS

František Pollák, Ján Dobrovič, Jan Váchal, Jarmila Straková, Petra Pártlová

Faculty of Corporate Strategy, Department of Management
Institute of Technology and Business in České Budějovice
frank.pollak@acuityeng.com

Keywords

Reputation management, Innovations, leadership, e-commerce, distribution, crisis management, global pandemic.

Abstract

The main aim of the paper is to present selected results of a comprehensive analysis of crisis management of corporate reputation of the best e-commerce entities operating on the Slovak internet market. On the sample of selected winners of the Heureka Group Shop of the year 2019 Quality Award poll, we conducted an in-depth analysis of their online reputation at the peak time of an ongoing pandemic. As part of the analysis, we examined the online environment from the perspective of all relevant determinants of reputation, the results were then clearly interpreted and offer a comprehensive view of crisis management of reputation in the online environment. The findings can serve as a basis for further research, or as a benchmark in the event of a recurrence of crises of a similar nature and magnitude in the future in order to eliminate shortcomings and maximize the effectiveness of marketing communication in brand building and brand protection.

1. Introduction

The Internet has changed the way we are thinking about reputation. What was once private is now public. What was once happening on the local level is now discussed on the global level. What was once ephemeral is now permanent. What was once trusted is now unreliable (Delina, Dráb, 2009, Delina, Tkáč, 2010, Delina, 2014). These changes happen because the Internet has modified our interaction with reputation (Dorčák, Štrach, Pollák, 2015, Pollák, Nastišin, Kakalejčík, 2015). Understanding the unique relationship between technology and online culture is a key to understand how to manage online reputation (Maryška, Doucek., Novotný, 2012, Maryška, Doucek, Kunstova, 2012, Loayza, 2013). Those who apply off-line techniques on their Internet reputation or use off-line assumptions to solve online problems are doomed to failure. Instead, the user must be capable to understand the cultural and technical differences between the Internet and off-line world to effectively protect and improve his online reputation (Fertik, Thomson, 2010). Walter (2012) argues that reputation is a cornerstone of one's life and business. This means that reputation is very fragile and one mistake can sometimes cause irreparable damage (Pollák, 2015). This is especially true in the digital world ruled by radical transparency and high standards of customers (Soviar, 2011, Soviar, Vodák, 2012). Entities must be able to learn to communicate on social networks, follow the "chatter"

on social media and effectively respond to such impulses without harming their reputation in line with expectations of their customers (Svetozarovová, Polláková, 2015). Siano et al. (2013) argues that when the Internet allows consumers to share information about businesses and brands, entities have the opportunity to control information published about them. Negative comments on the Internet can quickly and severely damage image and reputation of the brand. There is a wide portfolio of methods for quantifying reputation described in literature.

Speaking about reputation systems, the simplest solution is to sum up all the relevant positive and negative reviews. The total result related to the specific user is the difference between all positive and negative reviews. This principle is used mainly on eBay, one of the largest online markets and community with over 50 million registered users. After each transaction the buyer and the seller can give each other positive, negative, or neutral rating, which in turn adds plus or minus points (1, -1, 0) to their reputation. Users can also leave comments. When people leave negative rating, they usually leave a comment that explains it. Although the eBay reputation mechanism is very simple, empirical results show it supports transactions between sellers and buyers. It is mainly due to the fact that sellers with better reputation are more likely to sell more. Also, this mechanism can prevent people to artificially boost their reputation with each other (Resnick, Zeckhauser 2002).

Another important model for online reputation monitoring is ReGreT model introduced by Sabater and Sierra (2003). ReGreT model is a standard model of trust and reputation system aimed at a SME e-commerce environment where social relations between individuals play an important role. The system takes into account three different sources of information - direct experience, information from third parties, and social structures. ReGreT reputation model is based on three specialized types of reputation, first is proven reputation, calculated from the information coming from witnesses, second is reputation surroundings, calculated using information based on social relations between the partners, and finally third one, system reputation, based on roles and general characteristics.

Model Flow represents models that calculate reputation using transitive interactions of individual users. Some allocate the entire community (users' network) a certain amount of reputation which is gradually redistributed to all users. Thus, the reputation can be increased only at the expense of other users. The most famous example of this algorithm is PageRank from Google. Google has become the most powerful and the most popular Internet search engine. PageRank mechanism evaluates sites on a scale of 0-10. Each page starts from a scratch. Pages having 5 or 6 points are deemed very good. If the page has a score greater than 7, it is among the very big players. It is better to get one link from a Page Rank with a score 5 or 6 than have 20 links from pages with a score only 1 or 2 (Wang, Vassileva 2007).

Model TOR, created by us (Pollak 2015), takes into account n-variables, which we call reputators. Reputators can be freely defined as determinants of reputation, where each determinant must have a significant ability to influence the reputation of the entity in the online environment. The basic reputation in a number of variables is the percentage expression of the score from the extended sentiment analysis (Saxony 2015), where sentiments are gradually analyzed - the nature and polarity of search results in the top ten places in Google search for the analyzed subject. Other reputators are added to the equation on the basis of the object of research, by default it is various competitive scores of subjects created on the basis of the size of target audiences in the social media environment or the evaluation of relevant Internet players. The overall level of reputation then takes the form of the arithmetic average of the individual reputators. We apply this methodology for further analysis.

2. **Objectives and methods**

As mentioned in the introduction, the Internet has changed the rules of the game, the availability and authenticity of information plays a key role in customer decisions. Predicting customer behavior is

complicated even under standard conditions, however, what if a large part of the traditional B2C market suddenly collapses. How the absence or a significant reduction in the supply of brick and mortar players will complicate the already extremely difficult situation of managing an online reputation? On 16th of March 2020, a state of emergency was declared in Slovakia in connection with a pandemic caused by a new coronavirus. There was a partial closure of the economy and a significant slowdown, in significant causes lockdown of the brick and mortar B2C market. This hitherto unthinkable state created specific conditions in which it was possible to observe on a daily basis highly non-standard situations, which created an environment for research in almost all areas of scientific research. From the point of view of our study, it was the research of interactions and crisis communication of e-commerce entities operating on the Slovak market. It is the e-commerce market that has passed a stress test of unprecedented proportions. Almost overnight, this market had to replace the entire supply of traditional entities, which were forced to close their operations. The enormous pressure on the supply chain tested the stability of the system, problems of an operational nature (but) with strategic implications were solved in real time. The logistical flow of goods no longer had the traditional form of chaining, it was reduced to a minimum, couriers or dispensaries of goods took over a large part of the function of wholesale, retail and intermediaries. To this chaos we must also added the uncertainty nervousness of the market, and the effect of frontloading with supply shocks. The uncertain situation that created the cocktail of the high-explosive reputation bomb. In this situation, we carried out our research in order to identify basic qualitative data for furthermore comprehensive quantitative research.

The object of the research consisted of selected representatives of the e-commerce market. As it was difficult to directly select specific service providers while maintaining the necessary validity of the data, we decided to perform our analysis on selected ambassadors of a selected research group. Research group was represented by the sample of winners of the Quality Award of the Shoproku 2019 survey of the Hereka Group (Heureka 2019) in its selected six categories, which we evaluated as the most relevant given the current situation.

In the time interval 1st to 18th of April 2020, we monitored the market situation on a daily basis. It was a period when the disease was most likely culminated, and when, in our opinion, was already possible to record the first data generated by the crisis condition market situation. Data for further analysis were collected on the last day of our chosen time interval.

For analysis itself we used the TOR methodology (Pollák 2015), we chose the methodology based on its ability to identify changes in the individual determinants of reputation in a relatively short time.

In its first step we performed an extended analysis of the sentiment ASA (Pollák, 2015, Markovič, Dorčák, Pollák, 2017). ASA- extended form of sentiment analysis (SA) of the nature of first ten Google search results for chosen subjects, takes into account the occurrence of sentiments in several dimensions compared to the original methodology. In our case of our six subjects were tested for sentiment of search results both in the "general" and "news" categories of Google. Our aim was to find out how the uncertain situation affected the well-controlled internet presence of chosen subjects. From the point of view of ASA analysis itself, as a search phrase we used well-known and well-established name of the research subjects. As we mentioned above, the main factor in the process of ASA is the quantification of the sentiment of results displayed after typing key words to the search engine. The results may show positive, neutral, and negative feedback (see Table 1). These sentiments, in order words polarity direction of the text, as well as the position at which the result is displayed will give an idea about the research subject, thus ultimately determining its online reputation (Pollák, 2015). The process records the evaluation of the first 10 results in Google search. After summing up the sentiment points, we reach the final amount. That amount is then a starting factor in assessing the success or failure of companies in the particular segment. The following table (Table 1) shows the values assigned during the sentiment analysis:

Sentiment/ Position of the result	1	2	3	4	5	6	7	8	9	10
Positive sentiment (+)	20	19	18	17	16	15	14	13	12	11
Custom web site of the organization (x)	10	9	8	7	6	5	4	3	2	1
Neutral sentiment (±)	2	2	2	2	2	2	2	2	2	2
Negative sentiment (−)	−20	−19	−18	−17	−16	−15	−14	−13	−12	−11

Table 1: Sentiment Individual Results/Position of Results, Source: Roháľ and Sasko (2011) In: Pollák (2015).

The Table 1 shows chronological sequence of awarding points to the analyzed entities. Positive response or sentiment results in the increase of the score. The higher the position of this sentiment in the search result, the more points are awarded. Similarly, but with the opposite effect it/is works in identifying the negative sentiment. Points are deducted, the higher the position of the display, the bigger the deduction of points, and this significantly deteriorates reputation. For the two dimensions analysis, we perform the same procedures for each of them separately, in our case separately for the category "general" and separately for the category "news". The resulting amounts are calculated and converted to percentages. This is based on the assumption that within a single group the entity may receive a maximum score of 155 points – the ratio 1 point = 0.645%. For purpose of our analysis with 2 groups, the entity may receive a maximum score of 310 points – the ratio 1 point = 0.322%.

In the second step, we identified other relevant reputation determinants as follows:

- Google reviews score (converted to percentages),

- Facebook score (read as the nature of emoticons of interactions during the observed period / directly through the evaluation of the subject's profile, then converted to percentages),

- Customer rating score from the Heureka portal (the recommendation to realize the future purchase expressed by customers in the last 90 days).

In the third step, we proceeded to the calculation of the TOR indicator. Subsequently, we visualized and interpreted the findings, visualizations and interpretations are shown in the next part of the paper.

3. Results and discussion

The following table shows the partial and overall results of the analysis, represented by the values of individual reputators and the overall TOR indicator:

No.	Subject/ Result sentiment	ASA score (%)	FB score (%)	Heureka score (%)	Google score (%)	Number of pages indexed by Google	TOR score (%)
1.	DATART.sk	43.79	99.00	93.00	84.00	3,600,000	79.95
2.	Astratex.sk	70.52	92.00	98.00	0.00	578,000	65.13
3.	Footshop.sk	66.33	90.00	93.00	88.00	791,000	84.33

4.	MojaLekáren.sk	45.08	88.00	90.00	72.00	302,000	73.77
5.	svetnapojov.sk	37.67	98.00	99.00	0.00	77,900	58.67
6.	HEJ.sk	63.11	99.00	92.00	84.00	12,800,000	84.53

Table 2: Overall (Total) online reputation, Source: Own processing

As can be seen in the table, the ratings of individual reputators were diverse, in the case of the first entity, the giant from the traditional world, DATART we recorded in Google search results almost standard values for the model entity, in the first place were the entity's own pages, followed by pages with positive and neutral nature. Landing pages and affiliate-based sites have generated a fairly significant share of positive sentiment in search results. The near-ideal state within this parameter was not repeated in the "news" category search results. All search results were neutral and, in our view, were created by poorly optimizing content providers' sites, which added text from the contextual advertising window to messages of a different non-subject nature. From the point of view of the second subject in our sample, the Astratext.sk underwear store, in the category "news", it was an exemplary example of marketing communication through related media and PR news. The subject scored the full number of points in this category. Third in line, the Footshop.sk shoe store also demonstrated perfect PR mastery through media clearly targeting their customers. The fourth subject, the online pharmacy in the category of main search results achieved almost 50% of possible points with organic results supplemented by positive customer ratings, the better result reduced the gain of negative sentiments on the 9th and 10th ranks, having the nature of customer ratings. These could be eliminated thorough better work with the content. The category of "news" pointed to the need for a certain nature of PR, a significant part of the results had a neutral polarity, which greatly reduced the online communication potential of the subject. The online store with the alcohol- svetnapojov.sk achieved the lowest score in our comparison, which was mainly due to the total neglect of PR and marketing communication policy aimed at internet versions of important domestic media. The combination of key ones represented by the name of the subject generated only three mentions in the category of "news", two of which were in the nature of neutral sentiment. The last subject from our group, the shopping portal Hej.sk, showed a relatively solid media work both in terms of overall presence and in the case of PR and marketing communication, which we would recommend intensifying due to the size of the subject.

As an important finding we consider the fact that at the time of the research, the difficult pandemic situation was not transferred to the presence of subjects in form of their search results at all. This may be caused due to the longer response time of the Internet to shocks, we will continue to monitor this phenomenon.

From the point of view of the second step of the analysis, the review of reputators, it can be stated that these mostly organic indicators are evaluated highly positively by the subjects. Whether it is tens to hundreds of positive interactions with entities that do not have FB ratings, or relatively high ratings for entities that have them, to almost completely recommend customers for further purchase, these parameters only confirm the fact that they are the best players on the market. We only evaluate negatively the absence of Google's evaluation of two entities, which deprives them of their authenticity. However, this is probably due to the relatively short presence of the players on the market. Once again, we were surprised that the difficult market situation did not cause problems with reputation management during the period under review. In the analysis of hundreds of Facebook interactions in the research time period, we came across negative interactions only in trace amounts, if we take into account the nature of the market and the situation that generates stress and emotions in general, this is a very surprising and significantly positive finding.

In the following chart, we visualize the overall state of the reputation of entities, as well as state of their partial reputators:
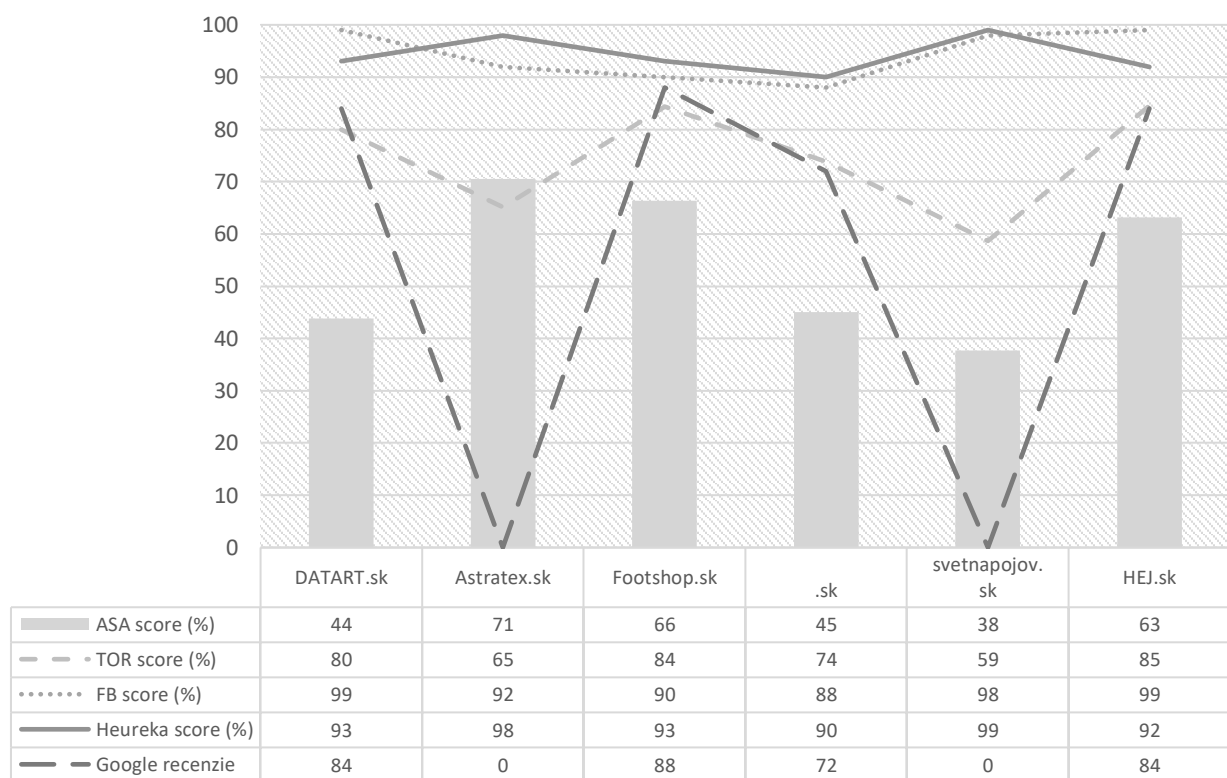
	DATART.sk	Astratex.sk	Footshop.sk	.sk	svetnapojov.sk	HEJ.sk
ASA score (%)	44	71	66	45	38	63
TOR score (%)	80	65	84	74	59	85
FB score (%)	99	92	90	88	98	99
Heureka score (%)	93	98	93	90	99	92
Google recenzie	84	0	88	72	0	84

Figure 1: reputation of entities, Source: Own processing

As already mentioned, we have identified another unusual phenomenon, which is a relatively high score of customer recommendations. Due to the difficult market situation, we consider it necessary to monitor this phenomenon over time and thus identify the possible methodological nature of the phenomenon, or the time response of the environment to change. In our opinion, both facts are worth exploring further.

In the third step, we calculated the overall TOR indicator, from the point of view of our sample, the highest value of the overall reputation was recorded by the e-shop HEJ.sk, followed by the e-shop Footshop.sk. With values of the overall reputation at the level of more than 80%, it can be stated that these are highly above-average values regardless of the object of research.

4. Conclusion

Specific conditions, in which we carried out our analysis created almost model situation for fulfillment our research aim. Investigation of interactions, crisis management and communication approaches of selected e-commerce entities provided prospects for future comprehensive quantitative research. E-commerce market has passed a stress test of unprecedented proportions. Almost overnight, this market had to replace the entire supply of traditional entities, which were forced to close their operations. The enormous pressure on the supply chain tested the stability of the system, problems of an operational nature, but with strategic implications, were solved in real time. Supply, distribution of goods to customers, customer service, these are all physical aspects of trading on the Internet. From the point of view of the sample we examined, we can conclude that the processes were handled well. Crisis management of the corporate reputation of selected e-commerce entities was appropriate to the situation. From the interactions observed on the social network Facebook during the period of research realization, we recorded no negative customer feedback. From the point of view of customer references, we recorded highly above-average satisfaction numbers. This

phenomenon represents the greatest challenge for further investigation, in order to identify the time lag of the impact of events on the reputation, respectively the methodological nature of the phenomenon. Both of these facts have a high research potential. From the point of view of the limitation of our research, we evaluate as the most important fact that we operated in the relatively limited time period, as well as we used relatively small sample for the data collection. Also, the processing and deeper interpretation of the results of our qualitative analysis will be the subject of subsequent in-depth research, as deeper research requires a higher allocation of resources, especially time. As our research was of a qualitative nature and aimed to reveal the context for further continuous research, we consider the limitations of the research as appropriate.

5. Acknowledgement

6. References

Delina, R. (2014). Transparency in Electronic Business Negotiations – Evidence Based Aanalysis. In: Quality Innovation Prosperity. Vol. 18, No 2, pp. 79-89

Delina, R., Dráb, R., (2009). Trust building on electronic marketplaces in the field of escrow services and online dispute resolution. In: IDIMT-2009 System and Humans-A Complex Relationship. Linz: Trauner, 2009. Pp. 149-156.

Delina, R., Tkáč, M., (2010). Trust Building Mechanisms for Electronic Business Networks and Their Relation to eSkills, World Academy of Science, Engineering and Technology 6 (71), pp. 380-390.

Dorčák, P., Markovič, P., Pollák, F. (2017). Multifactor analysis of online reputation as a tool for enhancing competitiveness of subjects from automotive industry. In: Ekonomický časopis. Vol 65, No. 2 (2017), pp. 173-186. ISSN 0013-3035

Dorčák, P., Štrach, P., Pollák, F. (2015) Analytical view of the perception of selected innovative approaches in marketing communications. In: Quality. Innovation. Prosperity=Kvalita. Inovácia. Prosperita. Vol. 19, No. 1 (2015), pp. 74-84.

Fertik, M., Thomson, D. (2010) Wild West 2.0: How to Protect and Restore Your Reputation on the Untamed Social Frontier. USA: Amacom.

Heureka Group. 2020. Cena kvality Shop roku 2019- Heureka Group a.s. Available from internet: https://www.shoproku.sk/vysledky

Loayza, J. (2013) The Beginner's Guide to Reputation Management: 8 Core Principles of Reputation Management. Available from internet: http://reputationhacks.com/guide-to-reputation-management-3-8-core-principles/

Maryška, M., Doucek, P., Novotný, O. (2012). Requirements of companies on the knowledge ICT specialists for the ICT administrator role. 4th World Conference on Educational Sciences. Barcelona, 2012. Procedia - Social and Behavioral Sciences. Vol. 46 (2012), pp. 4389 – 4393.

Maryška, M., Doucek, P., Kunstova, R. (2012). The Importance of ICT Sector and ICT university education for the economic development. 3rd international conference on new horizons in education – INTE 2012. Vol. 55, pp. 1060-1068.

Pollák, F. (2015). On-line reputačný manažment v podmienkach stredoeurópskeho virtuálneho trhu. Prešov. Bookman. (Slovak)

Pollák, F., Nastišin, Ľ., Kakalejčík, L. (2015) Analytical view on the use of mobile platforms in purchasing process. In European journal of science and theology. Vol. 11, No. 6.

Reichheld, F. (2011) The Ultimate Question 2.0 (Revised and Expanded Edition): How Net Promoter Companies Thrive in a Customer-Driven World. Harvard Business Review Press.

Resnick, P., Zeckhauser, R. (2002) Trust amnog strangers in internet transactions: Empirical analysis of eBay´s reputation system. Available from internet: http://presnick.people.si.umich.edu/papers/ebayNBER/RZNBERBodegaBay.pdf

Sabater, J., Sierra, C. (2003) Review on computational trust and reputation models. Available from internet: http://www.iiia.csic.es/~sierra/articles/2005/SabaterSierra.pdf

SAEC and Mediaresearch Slovakia (2013) Nakupovanie na internete 2013. Available from internet: http://www.nielsen-admosphere.sk/files/2014/08/nakupovanie-na-internete-2013-online-prieskum.pdf

Shoproku.sk (2016) Oficiálna internetová stránka najlepších E – shopov na Slovensku. Available from internet: https://www.shoproku.sk/vysledky (Slovak)

Siano, A. et al. (2011) Exploring the role of on-line consumer empowerment in reputation building:research questions and hypotheses. Available from internet: http://www.academia.edu/1096337/Exploring_the_role_of_on-line_consumer_empowerment_in_reputation_building_Research_questions_and_hypotheses

Soviar, J. (2011) Simplification of Marketing Scheme for Business Start-Ups. In Communications - Scientific Letters of the University of Zilina, Vol. 13, No. 4.

Soviar, J., Vodák, J. (2012) Value network as part of new trends in communication. In Communications – Scientific letters of the University of Žilina, Vol. 14, No. 2.

Stec M., Filip P., Grzebyk M., Pierścieniak A. (2014) Socio-economic development in EU member states - concept and classifications, Engineering Economics, Vol. 25, No. 5.

Svetozarovová, N., Polláková, T. (2015) Analytical approach to issue reputation management, its selected assessment and measurement tools, Global business conference 2015: proceedings : the possible positive aspects of economic crises. - Zagreb : Insitut za inovacije, 2015. pp. 228-234.

Svetozarovová, N., Polláková, T. (2015) Analysis performance appraisal methods of employees in terms of their impact on organizational efficiency. In Proceedings of the Third International Conference on Emerging Research Paradigms in Business and Social Science. Dubai, UAE: Middlesex University Dubai., pp. 130- 137

Walter, E. (2012) 10 Tips For Reputation And Crisis Management In The Digital World. Released on 12th November 2012. Available from internet: http://www.forbes.com/ sites/ekaterinawalter/2013/11/12/10-tips-for-reputation-and-crisis-management-in-the-digital-world/#2ca443c25075

Wang, Y., Vassileva. J. (2007) Toward trust and reputation based web service selection: A survey. [on-line]. [cit. 2014-02-09]. Available from internet: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.89.6365

BLOCKCHAIN IMPLEMENTATION
INTO SUPPLY CHAIN MANAGEMENT

Martin Potančok, Veronika Ptáčková, Tomáš Církovský

Faculty of Informatics and Statistics

University of Economics, Prague

martin.potancok@vse.cz, veronika.ptackova@vse.cz, cirkovsky1@gmail.com

Keywords

Abstract

The blockchain technology can be used in many ways in the pharmaceutical industry (e.g., fraud management, identity management, document verification). Blockchain can provide more insight and trustworthiness to supply chain processes in the pharmaceutical industry, where transport and storage conditions of medications are crucial. This article is based on a case study, a qualitative research method, investigating how the blockchain technology could be incorporated into supply chain management in the pharmaceutical environment. As blockchain stores and validates the entire data history, it provides the possibility to verify the authenticity of medications, check their handling, and monitor physical storage and transport conditions (such as, temperature and humidity) to prevent any unintentionally changes of their properties. The article proposes a design of an extended supply chain management process using blockchain. The main added value of using blockchain is having information about storage conditions during the entire lifecycle and better protection of the system against fake drugs.

1. Introduction

The focus of this article is supply chain processes in the pharmaceutical industry, especially tasks involving the verification of medication (drug) authenticity and any related operations. The spread of counterfeit drugs (Clark, 2015) and the Falsified Medicines Directive (European Commission, 2011) require changes to supply chain processes. According to Lee (2000), supply chain processes in the pharmaceutical environment are closely connected with information systems / information and communication technologies (IS/ICT). Blockchain is one of the most discussed modern technologies in today's IS/ICT industry.

According to Swan (2015), blockchain as a means of verifying transactions "can be used for any form of asset registry, inventory, and exchange, including every area of finance, economics, and money; hard assets (physical property); and intangible assets (votes, ideas, reputation, intention, health data, etc.)."

This article is based on a case study, a qualitative research method, investigating how the blockchain technology could be incorporated into supply chain management in the pharmaceutical environment. First, the general supply chain and the blockchain theories were analyzed on the basis of a literature review; second, an extended supply chain management process was designed in a case study and qualitative research. The structure of this article corresponds to the above. The main aim of this article

is to complement pharmaceutical supply chain management processes with the blockchain technology supported by IoT and geolocation. This article reflects the current situation in healthcare research and the effects of the continually expanding possibilities of IS/ICT.

2. Supply chain in the pharmaceutical industry

According to Schneller et al. (2006), supply chain management (SCM) in healthcare is described "by the information, goods and money essential to purchase and transfer the goods and services from the supplier to the final user in order to control costs". Schneller et al. (2006) deem SCM in healthcare to be more composite than in other industries. One of the reasons is an impact on human health (Beier, 1995). Patients require a wide range of products, medicines, equipment and comprehensive health services, which are very financially demanding for participants (Schneller et al., 2006). However, the diversity in the industry has positive effects in many fields, such as cost reduction, organization stability, and problem-solving (Mannix and Neale, 2005). In the pharmaceutical industry, not only economic conditions, but also the protection of human lives needs to be considered. It requires coordinating supply chains, planning deliveries between partners and having good knowledge of customers' and suppliers' needs. Quick access to relevant planning data, such as forecasts, production plans and inventory status data, is necessary (Jespersen, Larsen, 2005). Manufacturers monitor their supply chains very carefully because they need to understand the cost-to-serve for a diverse customer base (Shah, 2004). Shah (2004) adds that we are moving away from product diversification and local products. The European Union supports it, defining packaging and labelling standards.

Supply chains for pharmaceutical products is very similar to supply chains in healthcare. Inefficiency in supply chains is characterized by product discontinuity, product shortages, poor performance, and worse patient safety (WHO, 2008). Kaiser (2005) found that pharmacies are the final stop in pharmaceutical supply chains, as they are the last places before drugs are bought by patients/consumers. According to Kaiser (2005), manufacturers are at the beginning, followed by wholesale distributors, who buy drugs from them and distribute them to customers: pharmacies, hospitals, long-term care units and other healthcare facilities. Unfortunately, there are several issues related to pharmaceutical logistical processes, which have negative impacts on the cost and the quality of services (e.g., out-of-stock, shrinkage (Rossetti et al., 2008), high frequency of reorders (Danas et al., 2002) and counterfeit products (Schuster et al., 2007)). Brewer (2000) identified four indicators which are important for consumer satisfaction: improved quality (delivery on time, product and service quality, return and allowances, product durability), improved timeliness (source cycle time, supply chain response time and production plan achievement), improved flexibility (order handling, exceptional services, non-regular deals, order configurations, product variety), and improved value (service level, customer satisfaction). The next research will focus on a new type of supply chain optimization (Shah, 2004).

The pharmaceutical industry is an important sector. In 2018, the European drug market (pharmacies and hospitals at consumer price terms) was valued at USD 340 bn (EUR 301 bn) by BMI (2018). Unfortunately, digitalization is a problem in all sectors of the economy. As online pharmacies grow in popularity, many countries impose strict legal restrictions. The European Court of Justice permitted cross border sales of drugs in 2003 (European Commission, 2003). In 2011, EU Directive (2011/62/EU) set targets that EU/EEC Member States have to prevent the entry of falsified medicines into the legal distribution chain. Possible solutions include a systematic verification of security features at the point of issue and random verification by distributors. Medication verification repositories and other technological solutions are used in verification, however, not blockchain (European Commission, 2011). The potential risks, including fake drugs, meanwhile increase. The World Health Organization (WHO) characterizes fake (counterfeit) drugs as a silent epidemic. These pharmaceutical products can contribute to drug resistance and cause serious harm and death to

patients. According to a WHO survey, one million people on average die every year due to fake drugs. WHO (2008) warns against counterfeit and substandard medicines. The production of counterfeit drugs is big business nowadays and consequently a serious problem. For example, a half of all Viagra in online pharmacies is counterfeit. At the same time, it is impossible to calculate how many people buy counterfeit drugs. According to Clark (2015), some estimates suggest that it can be 5 million in the U.S. and around 2 million in the UK. The question is how to protect people from fake drugs.

3. Blockchain

Even though blockchain is not the primary focus of this article, it is necessary to provide a general description of the technology to facilitate better understanding of the following parts of this article. Blockchain was introduced in 2008 as a technology for the virtual cryptocurrency Bitcoin (Latifa et al., 2017), which is used by different types of users (Sudzina and Pavlicek, 2019). Blockchain is a decentralized database, where all transactions are stored. In the case of cryptocurrencies, a transaction is a movement of currency between accounts. All transactions are secured by digital signatures and a cryptographic hash function (Ryan and Donohue, 2017).

Blockchains consist of an initial block and transactional blocks. Each transactional block represents a single transaction and stores information about it, containing four main parts:

1. Data – depending on the specific field where the technology is used. In cryptocurrency transactions, it is information about the sender, the recipient and the amount of sent coins;

2. Timestamp – date and time of the transaction;

3. Hash – data compressed into a small string by a mathematical function;

4. Hash of the previous block – a reference to the hash of the previous block.

All transactions are cumulated and stored in a decentralized way in the entire blockchain system. As such, blockchains are decentralized secure tracking systems for data on every interaction. The blockchain technology can help with research and enhance transparency and patients' privacy (Benchoufi and Ravaud, 2017).
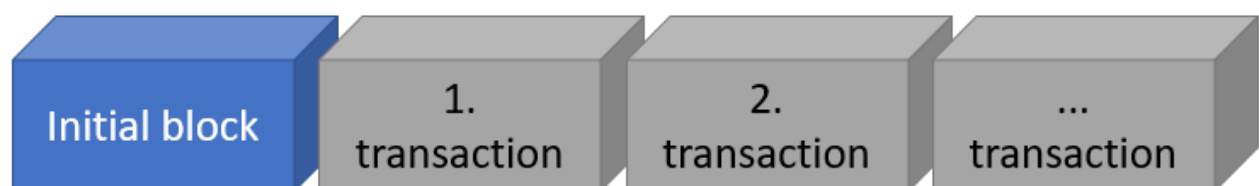
Fig.1. Visualization of blockchain transactions, Source: (Benchoufi and Ravaud, 2017)

From a private perspective there are four types of blockchains (Taskinsoy, 2019) – see Table 1.

Tab. 1. Main types of blockchains

Type		Read	Write	Commit	Example
Open	Public permissionless	Open to anyone	Anyone	Anyone	Bitcoin, Ethereum
Open	Public permissioned	Open to anyone	Authorized participants	All or subset of authorized participants	Supply chain ledger for a retail brand accessible to the public
Closed	Consortium	Restricted to an authorized set of participants	Authorized participants	All or subset of authorized participants	Multiple banks operating a shared ledger
Closed	Private permissioned "enterprise"	Fully private or restricted to a limited set of authorized nodes	Network operator only	Network operator only	External bank ledger shared between a parent company and its subsidiaries

3.1. Blockchain in pharmaceutical supply chain management

There are many areas in the pharmaceutical industry where blockchain can be used, for example, fraud prevention, identity management, document verification, etc. (Bocek et al., 2017). As stated above, this article focuses on the utilization of blockchain specifically in supply chain management in the pharmaceutical industry. The lifecycle of medications from the point of view of a pharmaceutical company has seven main stages:

1. Delivery of ordered material,
2. Storage of material,
3. Manufacture of an active substance,
4. Manufacture of final products and packaging,
5. Storage of finished products,
6. Delivery of products to healthcare organizations (see the pharmaceutical supply chain described above),
7. Selling/distribution of drugs to patients.

Blockchain and its principles can be used to monitor almost the entire lifecycle of a medication: from when materials are delivered to pharmaceutical companies for drugs manufacturing, to when medications are sold to patients. The main reasons for utilizing blockchain is to provide more insight and trustworthiness to supply chain processes in the pharmaceutical industry, where transport and storage conditions are crucial. The blockchain technology provides for data veracity and the tracking of transport and storage conditions:

1. **Data veracity:** As blockchain stores and validates the entire data history, it provides the possibility to verify the authenticity of medications and check their handling, when a code (e.g., a QR code, a barcode) attached to a medication is scanned. This can be very useful in tracking shipment statuses and fighting against fake drugs, which can seriously harm patients (as stated above) and cost pharmaceutical companies millions of dollars every year.

2. **Transport and storage conditions data:** As the blockchain technology stores all data, thus verifying medications origins, it enables us to monitor the physical conditions under

which medications are transported and stored, such as temperature and humidity, to prevent any unintentional changes to their properties (Guadelines for Warehousing Health Commodities). A violation of these conditions, such as a change in temperature, can induce chemical changes in a medication, and thus weaken its medicinal effect. Chemical changes can impact the activity, toxicity and stability of compounds (Giron, 2002).

The Bitcoin transactional model describes communication between subjects. As a container moves from point A to point B, two inputs and two outputs are created. One subject in the supply chain signs for it (an "I have the container" token) and then passes it on to the next subject, adding a transaction to the blockchain. The receiving subject confirms the delivery; and the process is repeated until the last subject. When the container reaches the final point, there is a complete, cryptographically verifiable, timestamped trail (Christidis and Devetsikiotis, 2016). Additionally, blockchain can support systems of early warning by providing data for monitoring proper storage and transport conditions and triggering alerts to stop a compromised shipment midway to save health and minimize costs.

All of the above-stated would not be possible without robust hardware and software, which must support enormous data flows integral to the solution. The following elements could be successfully utilized together with blockchain:

1. **Internet of Things (IoT)** – IoT should be considered to obtain data about physical conditions during the production, transportation and storage of medications. A set of sensors can be deployed in transport containers, machines and storage rooms. According to Zhang and Wen (2017), IoT and related E-business architecture has the advantage of systematization, high efficiency, flexibility, responsibility and low cost.

2. **Geolocation** – An accuracy rate and a success rate of deliveries are very important indicators in supply chain management of pharmaceutical companies. The indicators could be improved by placing GPS trackers into containers on transoceanic shipments. The tracking information could be included in the blockchain blocks for later verification and active tracking of current locations.

3. **ERP (Enterprise Resource Planning) systems** – ERP systems inform decision makers about customer orders, manufacturing plans, work-in-process, finished goods inventories, finance and accounting, enabling the sharing of information between departments (Mensah et al., 2015).

4. **PIS (Pharmacy Information Systems)** – Information systems deployed in pharmacies could verify data using blockchain in warehouse management, prescription dispensing, drug preparation and communication with health insurance companies (Anderson et al., 2009).

An extended supply chain management process taking into account the principles of the Falsified Medicine Directive (European Commission, 2011) and (Dvořáček, 2016) and integrating blockchain could be designed as shown in Figure 2.

1. **Delivery of ordered material** – Shipments are tracked by GPS and the physical conditions in containers are monitored by IoT sensors.

2. **Storage of material** – During storage, physical conditions are monitored by IoT sensors and all changes in stock are recorded in the blockchain database.

3. **Manufacture of active substance and final products, packaging** – The entire manufacturing process is monitored by IoT sensors in machines (this is already a part of existing pharmaceutical manufacturing processes). The only proposed change is to implement blockchain for further validation.

4. **Storage of finished products** – Physical conditions in which finished products are stored are monitored by IoT sensors in a warehouse.

5. **Delivery of products to healthcare organizations** – Shipments are tracked by GPS and the physical conditions in containers are monitored by IoT sensors.

6. **Selling/distribution of drugs to patients** – Thanks to blockchain, patients can verify if their drugs are authentic and if there were no issues during their manufacture, storage and transport.
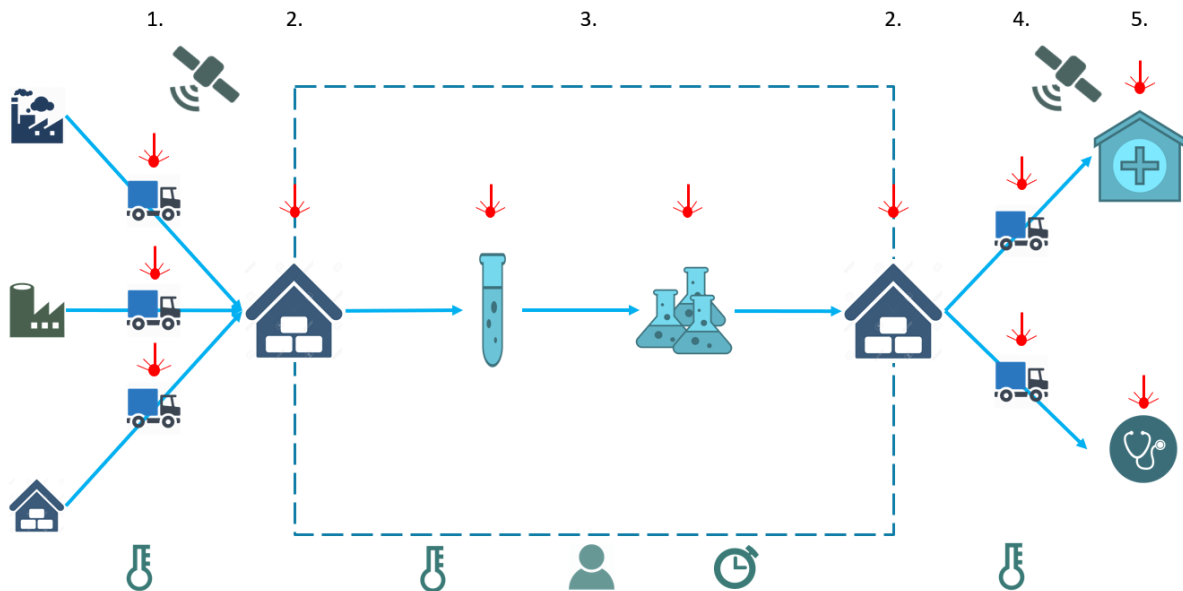

Fig. 2. Supply chain management using blockchain together with IoT and geolocation

4. Discussion

The proposed concept comes with a few challenges which need to be solved prior to large-scale implementation. The main challenge is the exchange of data during transport, especially over vast areas without signal, such as oceans, railroads in wilderness, deserts, etc., where continuous data flows are almost impossible. Possible solutions include using special containers with memories, where all data is stored during transport and uploaded to the blockchain database after reaching a connected destination; another possibility is sending data continuously (Sandoz, 2017).

The second challenge is decentralization, which is considered to be a cornerstone of the blockchain technology, as it provides for data trustworthiness and stability. In the case of cryptocurrencies, decentralization is achieved by utilizing volunteers, called miners, who are rewarded for operating the blockchain. With cryptocurrencies, anyone can be a miner, but this is not an option in the pharmaceutical industry mainly due to data security. However, big companies like IBM and Microsoft already provide platforms for blockchain, which can solve this challenge. Given the security needs of the pharmaceutical industry, Consortium and Enterprise seem to be logical choices out of the four blockchain types described above.

The last challenge is data security, because using blockchain technology throughout an entire supply chain process will generate enormous amounts of data containing sensitive information (such as, the types of products, their quantities and final destinations). This issue can be partially solved by introducing modern encryption technologies and best practices. Besides, blockchain can differentiate between owners and users of data, which could allow companies to concentrate on utilizing data without worrying about data security and compartment (Zyskind et al., 2015). This could also help

simplify regulatory frameworks and eliminate bureaucratic inefficiencies (Azaria et al., 2016). For example, Azaria et al. (2016) have proposed an innovation of electronic medical records based on blockchain. In this system patients have a comprehensive, immutable log and an easy access to their medical information in all medical institutions. The huge advantage is that the system uses the unique properties of blockchain.

5. Conclusions

This article proposes a design of an extended supply chain management process using blockchain. The main added value of using the blockchain technology is having access to information about storage conditions during the entire lifecycle of a medication and better protection of the system against fake drugs. The article confirms that it is possible to utilize blockchain specifically in healthcare, thus contributing to the growing body of research on this topic. As such, the article confirms the conclusions of Christidis (2016): *"The blockchain-IoT combination is powerful and can cause significant transformations across several industries, paving the way for new business models and novel, distributed applications"*.

6. References

Anderson PO, McGuinness SM, Bourne PE (2009). Pharmacy Informatics. USA: CRC Press.

Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). Medrec: Using blockchain for medical data access and permission management. In 2016 2nd International Conference on Open and Big Data (OBD) (pp. 25-30). IEEE.

Beier, F. J. (1995). The management of the supply chain for hospital pharmacies: a focus on inventory management practices. Journal of Business Logistics, 16(2), 153.

Benchoufi M, Ravaud P. (2018) Blockchain technology for improving clinical research quality. Trials 2017, 18:335.

BMI (2018). Pharmaceuticals & Healthcare Report.

Bocek, T., Rodrigues, B. B., Strasser, T., & Stiller, B. (2017). Blockchains everywhere-a use-case of blockchains in the pharma supply-chain. In 2017 IFIP/IEEE symposium on integrated network and service management (IM) (pp. 772-777). IEEE.

Brewer, P. C. (2000). Using the balanced scorecard to measure supply chain performance Peter C Brewer; Thomas WSpeh. Journal of Business logistics, 21(1), 75.

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the internet of things. Ieee Access, 4, 2292-2303.Clark, F. (2015). Rise in online pharmacies sees counterfeit drugs go global. The Lancet, 386(10001), 1327-1328.

Danas, K., Ketikidis, P., & Roudsari, A. (2002). A virtual hospital pharmacy inventory: An approach to support unexpected demand. Journal of Medical Marketing, 2(2), 125-129.

European Commission. (2018) A national prohibition on the sale of medicinal products by mail order is contrary to community law where it applies to non-prescription medicines which have been authorised for sale on the German market, Available at: https://curia.europa.eu/en/actu/communiques/cp03/aff/cp03113en.htm. Accessed 22 September 2018.

Dvořáček, J., (2016). Falsified Medicine Directive - ESM. [online]. 2016. Available from: http://www.sukl.cz/file/83549_1_1

European Union. (2011). Directive 2011/62/EU of the European Parliament and of the Council of 8 June 2011 Amending Directive 2001/83/EC on the Community Code Relating to Medicinal Products for Human Use, As Regards the Prevention of the Entry into the Legal Supply Chain of Falsified Medicinal Products. Official Journal of the European Union, 54(L174), 74-87.

Giron, D. (2002). Applications of thermal analysis and coupled techniques in pharmaceutical industry. Journal of thermal analysis and Calorimetry, 68(2), 335-357.B.D.

Jespersen, T.S. (2005). Supply chain management: in theory and practise. Copenhagen business school press, January, 11-14.

Kaiser (2005). Follow The Pill : Understanding the U . S . Commercial Pharmaceutical Supply Chain. Heal Strateg Consult LLC, 1–29.

Er-Rajy, L., El Kiram My, A., El Ghazouani, M., & Achbarou, O. (2017). Blockchain: Bitcoin wallet cryptography security, challenges and countermeasures. Journal of Internet Banking and Commerce, 22(3), 1-29.

Lee, H. L., & Whang, S. (2000). Information sharing in a supply chain. International journal of manufacturing technology and management, 1(1), 79-93.

Mannix, E., & Neale, M. A. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. Psychological science in the public interest, 6(2), 31-55.

Mensah, J., Annan, J., & Asamoah, D. (2015). Optimizing drug supply chain in hospital pharmacy department: an empirical evidence from a developing country. Available at: http://ir.knust.edu.gh/bitstream/123456789/11636/1/Optimizing%20Drug%20Supply%20 Chain.pdf. Accessed 22 September 2018.

Rossetti M, Marek D, Prabhu S, et al. (2008) Inventory Management Issues in Health Care Supply Chains. Available at: https://pdfs.semanticscholar.org/3aaf/c7c59614075672f6e6de42b0e76b57ba8b94.pdf. Accessed 23 September 2018.

Ryan, R., & Donohue, M. (2017). Securities on blockchain. Bus. Lawyer, 73, 85-108.Getting medicines to patients – the role of quality logistics (11 min read) | Sandoz. Sandoz, Available at: https://www.sandoz.com/stories/access-medicines/getting-medicines-patients-role-quality-logistics. Accessed 22 September 2018.

Schneller, E. S., & Smeltzer, L. R. (2006). Strategic management of the health care supply chain. Jossey-bass.Schuster, E. W., Allen, S. J., & Brock, D. L. (2007). Global RFID: the value of the EPCglobal network for supply chain management. Springer Science & Business Media.

Shah, N. (2004). Pharmaceutical supply chains: key issues and strategies for optimisation. Computers & chemical engineering, 28(6-7), 929-941.

Sudzina, F., Pavlicek, A. (2019). Impact of Personality Traits (BFI-2-XS) on Use of Cryptocurrencies. In: Hradec Economic Days. Hradec Králové, 05.02.2019 – 06.02.2019. Hradec Králové : Universita Hradec Králové, 2019, s. 363–369. ISBN 978-80-7435-736-7. eISSN 2464-6067. ISSN 2464-6059.

Swan, M. (2015). Blockchain: Blueprint for a new economy. " O'Reilly Media, Inc.".Taskinsoy, J. (2019). Blockchain: A Misunderstood Digital Revolution. Things You Need to Know about Blockchain. Things You Need to Know about Blockchain (October 8, 2019).

WHO (2008). Summary of the evidence on patient safety: implications for research. Available at: https://apps.who.int/iris/bitstream/handle/10665/43874/9789241596541_eng.pdf?sequence=1&isAllowed=y. Accessed 27 September 2018.

Zhang, Y., & Wen, J. (2017). The IoT electric business model: Using blockchain technology for the internet of things. Peer-to-Peer Networking and Applications, 10(4), 983-994.

Zyskind, G., & Nathan, O. (2015, May). Decentralizing privacy: Using blockchain to protect personal data. In 2015 IEEE Security and Privacy Workshops (pp. 180-184). IEEE.

CYBER SECURITY AND DIGITAL

IDENTITY AND PRIVACY

Michael Sonntag

Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
michael.sonntag@ins.jku,at

Keywords

Abstract

Security depends on permissions, which themselves depend on the identification of the person. This is contrary to another aim in IT security: privacy. This paper investigates the aspect of controllability of identification, i.e. whether the person to be identified can prevent others from doing this, and/or restrict the information disclosed to them to what the person provides her/himself. A method to "sign" stored messages, like forum posts, is presented, where authors can later on, for a self-determined time, prove identity (and therefore authorship) to interested parties and then disclose whatever information about them they desire.

1. Introduction

What is the "identity" of a user? Apart from the philosophical aspects of the question, this gets ever more complicated especially in IT. The reason is, that practically all security measures depend on first identifying the users and then (in various ways) assigning them permissions. However, the "securely and uniquely identifying a user" is getting more and more difficult and cumbersome. The past showed that e.g. passwords are problematic: requiring frequent changes leads to sticky notes, and remembering multiple complex passwords is too difficult for users. Password managers can help to a certain degree, but do not solve the basic problem, as this tool still has to – this time very – securely identify the user before it discloses the stored passwords. This leads to a reduction in the number of identification acts (making a complicated one more acceptable), but heightens their importance. An additional factor is the increasing use of mobile devices: they might contain a camera, 3D-scanner, fingerprint sensor etc, but these are often cheap and not very secure (see Roy/Memnon/Ross, 2017). Also, because of a lack of external interfaces e.g. two-factor identification with additional tokens is problematic (note that NFC might improve this, but such tokens do not seem to be in widespread use at the moment).

This is exacerbated with the mounting problems of biometric identification: deep fakes are now getting trivial to create. 2D face recognition should therefore be considered completely broken if unable to reliably detect them. 3D scanning might still be quite secure, but whether this remains so, is doubtful. Current subversion approaches require 3D masks (Thümmel/Sickert/Denzler 2019; Jia/Guo/Xu 2020) – but who can hide the shape of their head in public (and 3D printing is getting widely and cheaply available)? And faking the voice is today possible too. These approaches suffer from a problem of most biometric features: they are public. If they are not public, they typically cannot reasonably be used for identification purposes. For example, a picture of the genitals might not be easily available, but unlocking your phone in public would be impossible too. If a feature can be used in public for identification, third persons can use this public process (and often even without

315

the process!) to create a copy too. Some features might be more difficult to obtain or require a kind of trap (vein or retina scans), but still not completely impossible, and later to falsify for a specific person to be attacked deliberately. These increases in "hacking technology" lead to ever more invasive and complicated identification procedures – or a much lower certainty for identification, giving rise e.g. to multi-factor authentication

2. Identity and Privacy

If identification is getting ever more problematic, is this not a good direction regarding privacy? Unfortunately, it is not. While identification might not be as reliable any more, this applies against deliberate attacks to impersonate someone. On the contrary, for "normal" use it is getting better ("covert recognition": Prabhakar et al 2003). Face recognition might not be extremely reliable, but coupled with voice recognition and perhaps augmented by fingerprint scans and quick automated DNA testing (the last only a future prospect) persons can be identified in public easily. This also means they can be traced in their movements, contacts, interests… Also consider, that it is not always strictly necessary to identify the person itself – a mobile phone might differ from the identity of its owner, but it is a very good surrogate, as it is rarely given to others and almost continuously carried by the person.

These two aspects lead to a paradox: identification should be better, to improve security – but simultaneously worse, to enhance privacy. How can those two be combined – if this is possible at all? One option is to increase the *controllability* of identification, i.e. whether we *want* to be identified (→ quick and secure) or *not* (→ impossible to do it). This is quite obviously a good idea, but the problem lies in the implementation. We will therefore take a look at the various methods of identification from this point of view and how they support it.

2.1. Controllability of Biometrics

As discussed above, controllability of biometric identification is difficult. Everything that is publicly obtainable (viewable, observable) is hard to constrain. For example, gait recognition might be very secure and for humans almost impossible to replicate (Muaaz/Mayrhofer 2017), but if we can obtain exact measurements (carried device, long video), a robot could be easily built to exactly replicate these motions (or a software to replay these measurements as fake "sensor output"). This is true at least for all widely-used biometrics: they can be observed and be replicated.

- Fingerprints: Fingerprints are left everywhere, and there is absolutely no practical way to keep them secret. Also note that e.g. driving licenses or passports might require/produce large libraries of fingerprints, at least in the hands of the state.

- Face recognition (2D): Unless continuously wearing a mask (claims that this problem is solved, e.g. Pollard 2020, should be taken with a grain of salt – there simply is much less information available), there are just too many possibilities of obtaining a photo of your face. See also social media sites where a picture of any person might be uploaded and tagged, even though that person might not be a member of that platform at all.

- Face recognition (3D): Similar to 2D pictures, videos of a face from multiple angles are trivial to obtain from afar and without being noticed by the victim. Even deliberate and detailed IR scans (exactly like phones use) would hardly be detected by the victim.

- Iris scans: As has been demonstrated, a high-resolution picture from several meters away is enough (see also Fancourt et al 2005 for recognition via long distance). Therefore, this must be considered as impossible to keep secret too.

- Retina scans: Controllability exists only partially. The information might be obtainable, but only difficult (e.g. during an eye examination), especially from afar (flooding with infrared light, zoomed high-res video of eyes to obtain picture of whole retina might be possible).

- Hand geometry: If not permanently wearing gloves, the geometry of a hand can be deduced from pictures easily. Also, prints on surfaces can be helpful.

- Vein scans (hand, finger): This can be more difficult, as it requires infrared and the underside of the hand/fingertip. Still, if deliberately targeting a person, preparations can be made where a scanner is mounted beneath a surface where the person is going to put their hands.

- Gait recognition: The way of walking cannot be kept secret, as creating a video of someone walking is easily possible for as long as needed.

- Keystroke dynamics: Speeds, delay, etc of typing cannot be kept secret and might even be deduced from merely videotaping or listening to it (Asonov/Agrawal 2004).

- Heart pattern: Heart beat patterns have been monitored from afar with precision distance measurements. However, at the moment it is unclear whether this is enough to fool identification algorithms. But see also smart watches: if you hack/modify them, you definitely get all the data necessary, potentially even in real-time.

- Brain pattern: Probably at least at the moment it is impossible to detect brain waves at a distance, but verification is therefore similarly intrusive. This does not seem to be currently in practical use.

- Voice analysis: The voice is impossible to keep secret and it is today easy to synthesize arbitrary texts in a different voice.

- Signature recognition: Your manual signature is trivial to keep secret in theory, but easy to obtain if targeting a specific person, as they exist at many places. Advanced signature recognition not only takes the visual pattern into account, but also dynamics (speed, pressure); these are more difficult to skim, as only monitoring of an actual signing discloses these values (e.g. videos of a signature will not easily provide pressure information). Replication is probably easy with advanced technology readily available (simple robot with lots of degrees of freedom).

- DNA: There is no way to keep this secret at all. You distribute your DNA all over the area whenever being merely present.

Typical countermeasures against attacks on physiological biometrics (behavioral biometrics has this already built-in) are liveness checks. These usually can be countered relatively easily, as the degrees of freedom are limited, especially as the person still has to be recognized and therefore the biometric feature itself cannot change too much. This is more a question of "what exact feature is the verifier looking for – so what do I have to fake?" than impersonating a whole human being.

The best biometrics regarding controllability require strong physical closeness to obtain the data. The natural tendency of persons to keep a distance helps there. This is reduced with features that are regularly used with extreme closeness, i.e. by contact, like hands.

Controllability has a second dimension: are fakes easy to "present" under observation? For example, a 2D picture with the eyes cut out might fool lots of mobile phones into unlocking themselves, but if a human is observing the interaction, it will be practically impossible to not raise suspicion. A further aspect here is, whether the observer can detect manipulations, but does not receive identity information (e.g. using a fake external eye to fool a retina scanner is detectable, but neither in this manipulation attempt nor on real verifications the observer obtains the identity of the person

authenticating; but observing the entering of username and password discloses it). Similarly, while special makeup or clothing can prevent facial recognition by public video surveillance, this will easily be apparent and can probably also be trivially identified as "problematic" by an algorithm i.e., the aim for privacy would be not for the cameras to *not identify* the person, but *not detect the presence* of a person (which subsequently should be identified) at all.

The ideal biometric identification would be a kind of permanently closed eye that looks like normal skin (no recognition that it is there), that will solely be opened for identification (no chance of observing it during other activities) at very close range (no obtaining by observing an identification act), can be opened only deliberately (no accidental or prompted disclosure outside of identification) – and is not used for anything else (no possibilities for re-purposing "normal" actions). While this does not exist in biometrics, these features are very important for assessing identification in general.

2.2. Controllability of Possession

Possession typically means a token, which produces some sort of code (numeric, picture etc), either displayed visually or communicated directly to the other side (e.g. via cryptographic protocols). As long as the physical device is in your possession, controllability is typically good. But most of these tokens do not have any further security measures, so as soon as an attacker got hold of them, there exists no further controllability at all. Also, they might not be easy to recognize individually, i.e. swapping them with a similar looking device will only be discovered on an identification attempt. The only possibility is discovering the loss or trying a login and still possessing enough information to lockout the device for all entities that will attempt verifications in the future (note the possibility for DoS, so an alternative identification is required) – which is only possible if these potential verifiers are known or check at a specific instance for any updates of exclusions. While the secret contained within these tokens is most of the time not absolutely safe, i.e. with enough investment it might be possible to extract it, this takes a long time and changes/destroys the token. Therefore "undetectable copying", like in biometrics, is not possible absent security issues (see e.g. Charette 2011 regarding the RSA tokens, where probably the seeds were stolen).

Regarding privacy, tokens are usually very advantageous. They will be connected to a certain person, but this information does not have to be apparent from the device at all (unlike mobile phones). This means, stealing "some" device does not automatically disclose its owner's identity. Similarly, the act of using a token for logging in does not provide information for identification (separately from the rest of the login process, e.g. a username), as tokens are typically at least outwardly identical and do not use the identity information directly. However, the verification protocol might disclose the identity of the verifier to listeners. Moreover, even for tokens with NFC capabilities, many of them incorporate a physical button, meaning they cannot be incited to perform an unnoticed authentication without the knowledge of the owner. Therefore, large-scale identification of users for monitoring purposes is impossible with those kinds of tokens.

2.3. Controllability of Knowledge

Knowledge itself is controllable very well, as direct extraction from the brain is impossible. Extortion, violence, bribes etc are of course still options. However, as soon as the information has been disclosed, no further controllability exists at all, as there is no influence over any further dissemination. Additionally, as the knowledge originally only exists within humans, it *must* be disclosed in some way to serve as basis for identification. This act may be, and typically is, observable. Contrary to e.g. fingerprints it is usually easy to "copy" during this act too, as knowledge is almost always typing some kind of "code" into a version of a "keypad" (passwords, selecting symbols/pictures with a pointer, entering numbers on a keyboard, using hand or eye movements etc). Compared to possession therefore the act of identification is here a bigger problem of controllability.

Knowledge does not easily enable identification (absent an explicit act of identification) through mere observation; however e.g. "rare" passwords can disclose the identity through recognizing the person through it as the same one encountered somewhere else. The knowledge (or a matching counterpart) must be known to the verifier too, who additionally associates it to an identity. Privacy is therefore protected not very well. But while e.g. passwords are often not unique, knowledge elements are changeable and not tied to a person. i.e. someone can "change" their identity or "share" it with other persons or entities, e.g. software programs or agents, effectively creating a surrogate of themselves (which is impossible with biometrics and usually extremely hard with tokens).

2.4. **What is identifying you?**

Another aspect of "identification" is the question, what identifies you. E.g. an IP address might individualize you and allow certain person/institutions to discover your name, but your passport number – while much more reliable and unique – can often be more anonymous as fewer persons have access to it in ordinary business, or access to the special database storing them all. Essentially, "proving the identity" always takes place in relation to a *specific other entity*. That entity possesses certain knowledge, has access to some databases etc. The identity data disclosed could therefore be specifically tailored to the current counterpart to provide only them - and only just as much as necessary – that information. As long as that information is not passed on - and deleted soon after - this would work. But the knowledge of an entity may increase (or decrease) over time, and their data can be passed on to others with a different set of databases and more possibilities of deriving further information through combining them.

It is therefore better to take one step back and think about the *need* for identification. Mostly there is no actual need to know the unique identity of someone, merely to ascertain certain of their properties. For instance, when attempting to cross a border the identity of the person is of no interest to the border police: they want to ensure the person is allowed to cross the border, not on a list of fugitives/wanted person, old enough etc. This can be reduced to these classes of properties:

- (Not) Being on a list: e.g. persons granted access to a location or possessing a driving license. The exact position on the list as well as any other data (columns, date/time of being added to it etc) should remain unavailable. Examples for not being on a list are not being wanted or having had the driving license revoked. So sometimes multiple lists might have to be checked.

- Properties of the person, for instance age, being above a certain threshold age, gender, or social security insurance exists.

- Knowledge of some data, e.g. an access password.

- Possession of abilities or resources, e.g. a minimum of available computation power or a short response time (to prevent relaying).

- Ideally only a single element is disclosed in any identification process (e.g. "yes" or "no"), as then linking information, A+B combined with B+C produces A+B+C, is impossible

Note that there are different ways of achieving this. For instance, "being on the list" could be confirmed by a trusted third person (e.g. the issuer of the list), or a proof that a list provided by the verifier does include you (trivially: everyone on the list knows a shared secret value – proving the knowledge of the value is evidence of being on the list without disclosing the position on it). The latter is preferable from the point of privacy, as "confirmation by a third party" requires identification to that party. If unavoidable, these should be offline proofs (with all their disadvantages, like the revocation problem; but see e.g. OCSP stapling), so that no third parties need to be involved in the verification (and so at least know this fact and the time of it taking place).

3. Identification within synchronous communication

Identification within synchronous communication is very problematic, and getting increasingly so. For example, with deep fakes and voice faking the so-called "nephew trick" of impersonating a relative to obtain money, is going to be even more dangerous. And in times of a virus outbreak where old persons should not be visited in person – and simultaneously ever more old persons are tech-savvy and use electronic communication means – this could produce huge problems. If the person looks like the nephew (one photo from Facebook may be enough), sounds like the nephew (calling him under some pretense to get speech samples), and merely has a slightly different (but with a good explanation) online address, convincing someone to part with money or valuables can become much more likely and therefore dangerous. The methods for video identification, e.g. when remotely opening a bank account, are going to be either more complicated or unreliable, too. But at the same time, we can also pretend to be at multiple locations. E.g. video fakes coupled with a good chat-bot and a not-so-important-role, meetings do not necessarily have to be attended personally.

Consequently, how do we know the person on the other side of a call is who we think they are? This has always been a problem for strangers, i.e. people we do not know well. But now this is getting problematic even for familiar people. If this becomes a larger issue, demands for mandatory data retention or identification/registration/official communication accounts will increase and lead to a swift decline of privacy. Even then, capturing a Skype-/E-Mail/... account will only get more interesting for attackers, as these get more valuable and simultaneously harder to shut down and replace, as they become a single root of trust. Therefore, it will be necessary to strongly increase the identification before using them (additionally a very nice position for logging when used/position/time/device/service/…, but also communication partners/habits etc).

What is therefore needed is identification *inside* a synchronous communication (like video conference, chat, or audio call). This should be separate from the communication itself, so establishing the connection does not give any listener or intermediary (like the service provider) any (mostly unrealistic – but should be kept to a minimum; see below) information or proves the identity of the person using it. So end-to-end encryption is a necessary - but not sufficient - requirement. Then nobody outside the call can obtain provable identification, as the verification takes place only inside. This preserves privacy, as solely the actual communication partners securely obtain the identity of the other side, but requires this additional verification step. This does not completely solve the problem, as a connection still has to be established, i.e. some "number" has to be "dialed" to reach the intended communication partner by the initiator (the respondent does not necessarily need any information on the other side, which might remain anonymous to him or others). This "number" need not be a direct identification (could e.g. also be a tor hidden service), but the caller must be able to store it for later communication. Because of this it must be stable, i.e. not change over time (a lookup does not help – then the "number" merely transfers into the "name" to lookup). Because of this, privacy is necessarily reduced at least to the point of linkability: this communication is to the same person as the one in the past, even if we don't know exactly who is calling, respectively being called.

4. Identification of asynchronous communication

Asynchronous communication has the same problem: how to securely identify the origin of a message. This could of course be possible via the hosting service, i.e. the operator of the forum, blog, newsgroup etc. But regarding privacy that is not ideal, as that operator then has to securely store that identity – and verify it before. The same problem exists for electronic signatures – everyone can verify them – and in this way they disclose the identity to whatever degree is contained in the certificate. Therefore the possibility to post anonymously, but the author still being able to securely

identify themselves – if necessary – is desirable. Note that in this model it is completely up to the creator of the content to voluntarily disclose her/his identity, and to which degree. This means that a short (or longer) time after posting someone may try to verify the author of a message or obtain some other information (e.g. age, nationality, location) from them, but this cannot be enforced (fallback to anonymity) – only voluntarily granted, declined, or ignored. There exist several problems that need to be solved when implementing such an approach:

- The message must contain some contact information, but this information may neither allow to identify the author, nor provide linkability, i.e. the possibility to declare that this message must be from the same person as arbitrary others. Anonymity also extends to the process of identification: starting the process should not reveal any information about the identity of the author. Ideally the author can stop the process of identification at any point before the successful end without giving away *any* information.

- The hoster/operator of the forum cannot identify the author from any stored information, i.e. the posting is anonymous or all tracing information is removed. This is not absolute, as e.g. the identification feature could also be used to securely prove the identity – and so the permission to post - as opposed to some claimed identity (service operators often identify users merely via an E-Mail address).

- The process of identification should not involve the hosting service/web site, i.e. they should know that someone looked at the posting, but not whether this person tried to identify the author or not - or what the author's identity is if the process completed successfully (and whether it did so). Of course they may attempt the identification themselves, but merely like any third party.

- The author should be able to arbitrarily decline the request. This could be e.g. based on elapsed time ("will prove for 1 month only, then it should have been deleted anyway") or conditional to the asking person willing and able to prove their identity (and being "acceptable" according to an arbitrary decision by the author). In case of acceptance, any kind of identification should be possible (see above; e.g. not restricted to a certificate).

- As no central registry of identities exists, collisions in the contact information of an article may occur. In case there are several posts with identical data, multiple persons must be able to be contacted. It must always be possible to reach the "correct" person for identification, regardless how many candidates exist - as long as this person is willing to identify. This simultaneously produces plausible deniability.

The system described here supports identification in synchronous calls too via communicating a "verification code" that allows the other side to initiate an identity verification.

A possible implementation for this problem could consist of the following elements:

1. To ensure asynchronicity, it cannot depend on the author being logged in, reachable at the moment etc. So a surrogate of the user in the form of a software agent is needed (or some kind of "message store" like an E-Mail account). This can execute and remain active around the clock. As long as no authentication takes place, it needs very few resources, as there is nothing to do for it. If this is not needed (e.g. for synchronous communication), some software is still needed, but it need not be active without the user.

2. When a new message is authored, the person tells its agent to "sign" it. The agent then creates a new public/private keypair and uses this to create a Tor hidden service. As this is separate for every signature, i.e. post, no linkability exists through them (if desired, the same key can be used for all posts on a single site or based on any other distinction too, enabling limited linkability). Simultaneously it prevents identification of or recognizing the agent e.g. through

its IP address (note that posting should then be done via Tor too, or this information – if stored by the hoster – potentially allows identification). This hidden service is the "entrance" point for verification. Whenever the author intends to "abandon" the post temporarily or permanently, the hidden service is simply deregistered. Verification will then always fail for lack of a responder to the inquiry.

3. The agents signs the message with an additional secret value (which could be the same for all messages or again individually – here this would only improve security). The exact algorithm depends on the kind of verification attempted later (see next element). This is separate from the private key for the hidden service.

4. If asked to verify, the agent uses a Zero-Knowledge proof to prove to the inquirer (which might have to identify themselves before) that it indeed knows this secret value. It then provides any identification information it deems appropriate for this verifier. As this is a Zero-Knowledge proof, the inquiring person/agent can rely (to an arbitrary degree) that this response comes from the real author's agent, but third parties cannot (if both collude, they could falsify the protocol in a way undetectable, and therefore still believable, to such third parties). For third parties it therefore looks like someone proved authorship of a post, presented an identity, and someone else believed this. They cannot verify it themselves, but can reasonably depend on the result, as most verifiers will be independent and not lie. Note that third parties can obtain the identification data only from the verifier, not from listening in on the (encrypted) protocol.

Such a Zero-Knowledge proof can consist e.g. of the discrete logarithm (as described by Chaum/Evertse/van de Graaf 2000): $g^x \mod p = y$. The message (or its hash value) is used as the base g and the secret number (perhaps randomly generated for each post or always identical for this user) as the exponent x. The modulo number p can be publicly known or, similar to the result y, be added as the authentication data to the post (in addition to the Tor hidden service descriptor). The agent can then prove the knowledge of this x without actually disclosing it by (repeatedly) correctly answering one of two possible challenges by the verifier. As the message is also part of the equation this will only work if it is unmodified, effectively creating a "signature". As the secret value is never disclosed, even a successful verification does not allow a malicious verifier to impersonate the author later on or convince third parties that the identification is correct (or wrong). Afterwards, any identification data desirable can be transmitted.

There remains only the problem identified above: as each agent individually and randomly generates a new hidden service identifier for each post, collisions may occur. While these are extremely unlikely (Tor hidden services are essentially public-private key pairs and a collision means that two agents accidentally generated exactly the same keys), they cannot be ruled out. Additionally, this possibility even has some advantages, as an agent can then plausibly decline that this is not "his" hidden service, even if it should be proven that he employs it for *some* message. As long as it cannot be forced to show this alternative message, tracing a hidden service only gives a (very) high probability that this agent is responsible for it. But as the Tor system is designed for only a single hidden service behind every descriptor, if two servers exist with the same, usually the later one to register wins. However, there seems to be no quick deterministic result – some clients could still see (at least for some time) the old descriptor. To correctly support the approach described here, an "enumeration" of all servers would be necessary, i.e. connecting to the first, if this one declines/fails the proof connecting to the second and so on. Load balancers for hidden servers do exists, but these work by knowing all the servers they are balancing – something that cannot be assumed in this context. This problem is therefore currently unsolved; a solution would need a change to the directory of hidden services.

Controllability of this approach rests in the hands of the author: she can instruct the agent to abandon individual posts, or simply destroy the agent, thereby rendering everything anonymous. A further advantage of this approach is, that even later advances in cryptography will not help. Should the secret for the Zero-Knowledge proof be discovered, other persons can claim authorship too, but this does not allow anyone to identify the author or proof such a suspicion. Similarly breaking the hidden service key would only allow impersonation, but not attributing it to a specific agent. Privacy is practically complete, as apart from any registration/login requirements to be able to communicate, no identifying information about the author is disclosed at all and there is no way to link it to another communication act.

The biggest drawback is the mirror image of the advantages: the author of a post cannot be discovered by anyone, not even the police. If a specific suspicion for a certain person exists, first the agent of that person would have to be found/identified. This should normally be possible (note that it could also be created/contacted/… via Tor, rendering this at a minimum extremely difficult). After analyzing the agent, it can be proven that the secret key for the hidden service is present, and that the secret for the Zero-Knowledge proof is available in it too. In effect the agent can be "forced" (or a similar agent provided with the necessary data) to perform the validation of the signature. As the agent must be able to provide the identity to be useful in some form, it has to be accessible to it too – unless that data is encrypted with a password known only by its owner, but then no automatic identification is possible. Optionally only an encrypted version of the secret data could be stored, with the matching key being part of the post – only this post can then unlock at most that data, effectively predetermining the data to be potentially disclosed and preventing extraction of the key without knowledge of the post (not very useful, as the message is typically known and otherwise can often be discovered through the hidden service public key which is present in the agent and part of the message too).

Essentially, verifying the authorship of an already known person with a known post remains possible (to a very high certainty – see hidden service collision) and discloses at least all the identity information stored in the agent for this post. As technically the verification is easy, testing all the contained data against a large number of posts is feasible too, as is searching for posts with a certain hidden service descriptor. Securing the agent against unauthorized access to its data is therefore paramount. It should therefore be executed only on a trustworthy server or encrypt the zero-knowledge proof data – which then requires the author to individually grant each identification process by providing the information necessary to decrypt it.

The only countermeasure if the author does not provide her identity or refuses to participate in identification is deleting the post (terminating the phone/video call etc) – or trying to get hold of the (respectively all existing) agent somehow and then try extracting the verification data as explained above. It should be noted, however, that at the moment the situation is very much the same: obtaining an E-Mail address via Tor might be getting ever more difficult (e.g. Google requires verification via a telephone number), but remains possible. Registering for most platforms only requires an E-Mail address and no further proof of identity. Therefore all messages are anonymous, although potentially traceable (IP addresses in logs, ISP records of IP assignments etc). But contrary to the scheme presented here, proof of authorship is very hard or impossible (especially without giving away information about other posts or platforms), as only current (i.e. not necessarily when the act took place) control of the E-Mail address, which is still some way from an identification, can potentially be proven. A digital signature, which is of course possible now, does not provide any controllability, as everyone can verify it at any time.

Compared to the elements for ideally controllable identification: In the newest versions of Tor it is no longer possible to enumerate hidden services, therefore the existence of an agent is not apparent (no recognition that it is there). Contacting such an agent is used solely for identification (no chance

of observing it during other activities). Observing the identification does not disclose the identification element, as a Zero-Knowledge proof is used. The agent's software or the protocol could however be buggy, but this can be verified quite well, so accidental or prompted disclosure outside of identification is very unlikely. As the protocol is solely used for identification, the agent cannot be fooled into thinking this is a "normal" interaction (no possibilities for re-purposing "normal" actions). This system therefore almost completely fulfills these requirements.

5. Summary

A method to prove the identity of a person by generating a privacy-friendly signature was presented and its implementation discussed. It can also be used in synchronous communication, where the implementation through an independent software agent would allow it to take place automatically in the background, but only with additional software on both sides (not performable by humans).

Generally, identification is the natural enemy of privacy – which no longer applies if nobody can identify the entity with reasonable means – it is therefore very important to put identification into the hands of the person itself - it should be impossible without their consent (controllability). Especially in time with a virus epidemic and increasing demands for smartphone Apps logging all interactions/meetings between persons, restrictions on identification by others become supremely important.

6. References

Asonov, D., Agrawal, R. (2004). Keyboard Acoutic Emanations. IEEE Symposium on Security and Privacy, 2004. Berkeley, 2004, 3-11.

Charette, R. (2011). SecurID - Used by 30,000 Organizations and 40 Million People - Sees Security Reliability Partially Compromised. IEEE Spectrum, 18.3.2011, https://spectrum.ieee.org/riskfactor/telecom/internet/securid-used-by-30000-organizations-and-40-million-people-sees-reliability-partially-compromised

Chaum, D., Evertse, J.-H., van de Graaf, J. (2000). An Improved Protocol for Demonstrating Possession of Discrete Logarithms and Some Generalizations. Proceedings of Advances in Cryptology — EUROCRYPT' 87, Springer, Berlin, Heidelberg, 2000, ISBN 978-3-540-39118-0, 127-141

Fancourt, C., Bogoni, L., Hanna, K., Guo, Y., Wildes, R., Takahashi, N., Jain, U. (2005). Iris Recognition at a Distance. Audio- and Video-Based Biometric Person Authentication, LNCS 3546, 2005, ISBN 978-3-540-27887-0, 1-13

Jia, S., Guo, G., Xu, Z., A survey on 3D mask presentation attack detection and countermeasures, Pattern Recognition, Vol 98, 2/2020, Elsevier

Muaaz, M., Mayrhofer R., Smartphone-based Gait Recognition: From Authentication to Imitation. IEEE Transactions on Mobile Computing, Vol. 16, No. 11, 3209-3221, 2017

Pollard, M., Even mask-wearers can be ID'd, China facial recognition firm says, Reuters 9.3.2020, https://www.reuters.com/article/us-health-coronavirus-facial-recognition/even-mask-wearers-can-be-idd-china-facial-recognition-firm-says-idUSKBN20W0WL

Prabhakar, S., Pankanti, S., Jain, A. K., Biometric recognition: security and privacy concerns, IEEE Security & Privacy, vol. 1, no. 2, 33-42

Roy, A., Memon, N., Ross, A. (2017). MasterPrint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems, IEEE Transactions on Information Forensics and Security, 12(9), 2013-2025

Thümmel, M., Sickert, S., Denzler, J., Facial Behavior Analysis using 4D Curvature Statistics for Presentation Attack Detection, Submitted for publication in IEEE International Conference on Automatic Face & Gesture Recognition (FG), arXiv:1910.06056 [cs.CV]

SECURITY OF IOT DEVICES BASED ON LTE

Nikola Židková, Miloš Maryška

University of Economics, Prague, Czech Republic
nikola.zidkova@vse.cz

Keywords

Abstract

The IoT (Internet of Things) is a system of connected devices which are usually called as a smart device. Now IoT devices are an important component of everyday life and their importance is increasing every day. An important factor related to the increasing amount of IoT devices is an increasing amount of data generated by these devices and their processing. New threats are identified almost every day, and these are making security and credibility more and more important. This paper is devoted to the security of IoT devices based on Long Term Evolution (called LTE) technology. Paper introduces the selected communication technologies and their different security solutions. This paper also identifies security threats based on the EBIOS methodology. The aim of this paper is to identify threats when using LTE technology. We are summarizing recommendations that are focused on common users of IoT devices based on the analysis of all significant security aspects based on EBIOS methodology. One of the key findings of this paper is confirmation, that the users themselves can significantly contribute to the security of IoT devices and thus increase the security and credibility of the entire IoT system.

1. Introduction

Internet of Things (IoT) is a dynamically growing part industry. The IoT term is used mainly for devices that are collecting and storing data and especially providing defined functionality (e.g., managing heating in intelligent houses, managing whole buildings, remote accessing to car data).

New IoT devices appear every month. Usually, these devices can be can connect to networks through one or more connection methods (IOT Now, 2019). We distinguish two key connection methods: wired cable and wireless connection. Both have their positives and negatives. A wireless connection offers some advantages over a conventional cable connection, such as the free movement of a device. However, such benefits are especially with security risks.

The most frequently used technologies are Bluetooth, 2G/EDGE, 3G-GPS/GPRS, Cellular 4G/LTE, Wi-Fi, Zig-Bee, Z-Wawe, 9LowPAN in the Czech Republic. The most used communication technologies in IoT devices are LTE, Wi-Fi, Bluetooth and ZigBee (RS Component, 2015).

At present, the literature doesn't examine comprehensively the technical and security aspects and issues of the interconnection of IoT technologies, and therefore this article focuses on diagnosing the risks associated with LTE technology as one of the ways of communication between IoT devices. The aim of the article is to provide a security risk analysis of Long-Term Evolution – LTE technology as a representative of the Wide Area Network (WAN).

2. State of the Art

The main problem of IoT definition is that they are not so much a definition as a vision. One of these visions says that the basic prerequisite and goal of the Internet of Things is to connect unconnected. Based on this "definition" we can say that devices that are not currently connected to a computer network will be connected. This connection will allow them to communicate and interact with other devices/objects or people. IoT devices allow us to control the physical world by making devices smarter and connected through a smart network (Hanes, 2017). Literature is also saying, that IoT consists of devices with communication capabilities, computing power, and local decision-making in a limited context. Communication can take place via any wireless or wired mechanism. However, wireless methods are typically preferred as they eliminate wiring costs (Sinha & Park, 2017).

The main factor why IoT is such a major security challenge is its own extremely fast growth. Manufacturers and developers are inventing and developing IoT devices without sufficient time for testing and with the lowest possible price. Safety precautions often go aside (Keenan, 2017).

A secure IoT solution includes several levels that combine important IoT security features in four different layers: the device itself (A), communications (B), cloud (C), and lifecycle management (D). These levels are graphically illustrated in Figure.
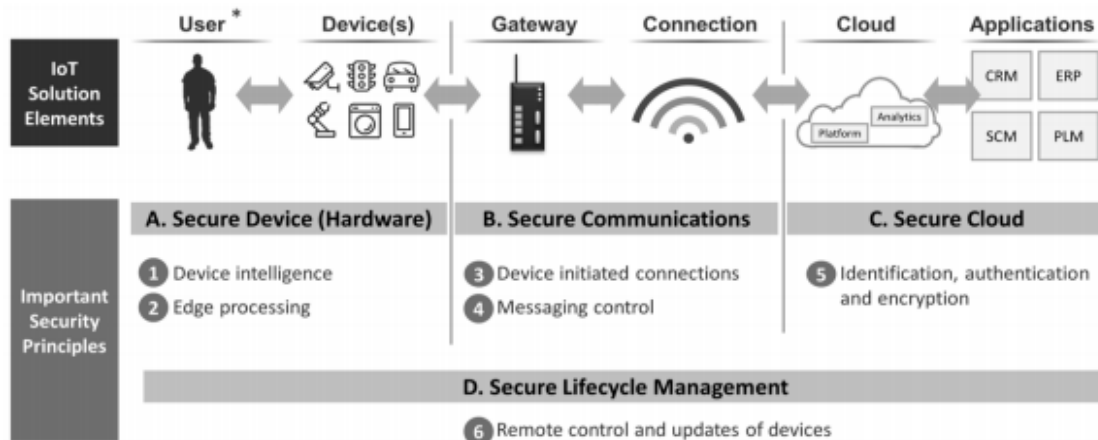
Figure 1. IoT security layers, source: (Padraig, 2016)

In this paper, we focus on layer B (part of the connection), namely in the area of LTE.

2.1. Network Architecture

Network architecture is based on 4 components:

- User Equipment (UE) are end devices in the architecture.

- UMTS Terrestrial Radio Access Network called E-UTRAN is a network consisting of base stations. The Base station is called evolved node B (LNB) in LTE. This eNB modulates and demodulates radio signals for communication with the UE. Said eNB then acts as an electrical switch point for Create and send IP packets to and from the primary network.

- Evolved Packet Core - EPC provides routing, and at the same time, it acts as a computational brain for the LTE network. EPC together with LTE forms the Evolved Packet System (EPS), which ensures the connection of the device to the IP network.

- Protocol network - The LTE network consists of a UE, an eNB group and another radio

- the access device constituting the E-UTRAN and the backend primary network.
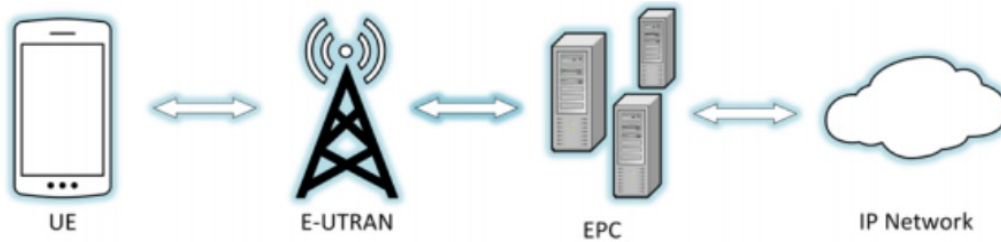
Figure 2. Cellular Network, source: (Cichonski et al., 2017)

2.2. Broad Band Generations and Standards

LTE is a mobile broadband communication standard whose networks are deployed worldwide. LTE belongs to the 3G / 4G mobile generation.

The first generation (1G) was an analog technology that had existed since the 1980s. In the 1990s, the first digital systems in the second generation (2G) were deployed. Voice messages, SMS and data services have been introduced. 2G technologies include GSM / GPRS and EDGE (Enhanced Data rates for GSM Evolution). 3G systems are based on GSM and radio access technologies, which brought the expansion of data rates for EDGE, as well as the further development of UMTS and High-Speed Packet data Access (HSPA). LTE and LTE-Advanced then crossed the generation line, offering the next generation of capabilities. The introduction of 5G will be the result of improvements in LTE, LTE-Advanced and LTE Pro (Cichonski et al., 2017).

Compared to previous 2G and 3G networks, LTE is roughly ten times faster in both directions, with up to 100 Mbps and up to 50 Mbps uploads. Works on frequencies in bands 800, 1800 and 2100 MHz. Bands above 1000 MHz are used mainly in highly populated areas.

2.3. LTE security

Within LTE security, authentication, cryptographic protection mechanisms, hardware protection, and LTE network protection mechanisms. LTE introduced a new set cryptographic data and key structure different from previous generations.

There are available 3 sets of cryptographic algorithms for confidentiality and integrity referred to as encryption algorithms EPS Encryption Algorithm (EEA) and EPS Integrity Algorithms (EIA). While these new algorithms were introduced in LTE, networking implementations usually include older backward compatibility algorithms with older devices. Many keys in LTE are 256 bits long, but in some current implementations, only 128 bits are used.

Hardware security is provided by UICC. UICC is a new generation SIM card on modern mobile devices and is the foundation of the LTE security architecture.

The primary LTE authentication mechanism of the device is the AKA protocol (Authentication and Key Agreement). Using AKA in LTE requires a document from 3GPP named TS 33.401. The AKA protocol is used for mutual authentication of subscriber/network. (Hegerová & Dostálek, Libor, 2016)

Air interface security refers to UEs and eNBs that communicate via a radio frequency link referred to as a Uu interface. These radio waves are transmitted from the UE antenna over the air until they reach the eNB antenna. This wireless communication is not necessarily private, which means that anything within the wave path can pick up these radio waves.

3GPP has specified optional features to protect the confidentiality of various network interfaces LTE. There is optional protection between the UE and the eNB on the Uu interface and the communication

between the eNB on the X2 interface secrecy. TS 33.401 specifies that the use of Internet Protocol Security (IPsec) in accordance with TS 33.2104 NDS (Network Domain Security)/IP should be implemented to provide confidentiality on the S1 interface, but the specification further notes that if it is S1 interface trusted or physically protected (Cichonski et al., 2017).

3. Methodology

Basic methods of data and information collection using secondary sources analysis and document analysis are used for this paper.

The main source of data are reports focused on security and threat analysis published every year by ENISA (European Union Agency for Cybersecurity). The ENISA report contains a summary of the key cyber-threats. This paper used ENISA reports for the years 2015-2019. Every report provides an overall assessment of situations and activities in the are of cyber threats in selected domains. Reports are based on data provided by MISP (MISP - Open Source Threat Intelligence Platform & Open Standards for Threat Information Sharing) and by CYjAX company.

The EBIOS methodology (Expression of Besoins and Identification of Objectives of Sécurité - Expression of Needs and Identification of Security Objectives) (EBIOS, n.d.) has been selected for the processing of risks related to communication technologies.

EBIOS methodology can be divided into 4 steps. First one is definition of the context and parametrs that are considered in the analysis of risks related to cyber. Second one is analysis of security needs and third one is analysis of potential sources of threats and thus od dreaded events. The fourth step is a risk assessment in terms of severity and probability. The second and third steps take place simultaneously. This step identifies three categories of threat sources and lists the common threats.

For each attack analyzed in the EBIOS methodology are defined main key parameters. Those are levels of impact severity, threat risk, and interest categories are added. Detailed description of each parameter can be found in Židková et al., 2020.

4. Results & Discussion

Cybercriminals' interest in IoT devices is constantly growing. In other words, it means that the greatest threat to IoT is a man. The purpose may vary, some attackers may want to gain unwanted access to a selected person, or by intercepting wireless devices, and some attackers may want to obtain confidential information. This paper focuses on the threat to IoT through one LTE.

4.1. Basic analysis of LTE security risks

The threat survey is based on an analysis of the 3GPP Working Group. The threats have been identified through academic research. Some of these threats may impact availability and network resilience, others are limited to the integrity and confidentiality of user data. Most threats would also affect only a limited part of the network. With the increased availability of cheap hardware and LTE integration software, many threats can be implemented at a low-level complexity.

General LTE threats are malware attacks related to end-user devices, malware attacks related to the infrastructure of Radio Access Network, malware attacks related to core network infrastructure and unauthorized access to operating and access network.

Key threats affecting LTE security are listed in Table 1 below.

Table 1. Definition of LTE threats, source: (Frankel et al., 2007; Mohamad Noor & Hassan, 2018)

ID (Risk)	Threat	Examples
R1	Botnet	Prevention of emergency telephone calls.
R2	Exploit kit	Unauthorized access to Operational and Access Management.
R3	Malware	Malware attacks on user equipment, Radio Access Network infrastructure, basic infrastructure.
R4	Web Application Attacks	Theft of services.
R5	Cyber Espionage	Device and Identity Tracking, Air Eavesdropping interface.
R6	Denial of Service	Radio Interference Attacks; sending unverified messages (ATTACH REJECT) leading to DoS; physical Attacks on Network Infrastructure.
R7	Web Based Attack	Downgrade; Pre-Shared Key (PSK) Attacks.

To further analyse the security of LTE networks for use in IoT, the identified threats had to be further processed and assigned one of the threat severities levels (see section 3). The mapping of LTE threats and their impacts at the threat severity level is performed in Table 2.

Table 2. Severity of impact for individual LTE threats, source: authors, data: (Kidston et al., n.d.; Qiu et al., 2006) and risk and categories, source: authors, data: (ENISA, 2016, 2017, 2018, 2019) and (Židková et al., 2020)

ID	Threat	Severity of impact	Risk	Attack category
R1	Botnet	Critical	Strong	Process Control
R2	Expoit kit	Critical	Strong	Process Control
R3	Malware	Critical	Significant	Process Control
R4	Web Application Attacks	Nebliable	Strong	Process Control
R5	Cyber Espionage	Restrictive	Significant	Spying and stealing data
R6	Denial of Service	Critical	Maximal	Process disruption
R7	Web Based Attack	Significant	Maximal	Process disruption

No device can be secure so high to avoid every possible attack. Developers and users of IoT devices can implement security policies and implementation of security mechanisms which can reduce potential risks. Risks for selected threats are shown as well in Table 2. The risk of an attack is calculated based on the occurrence of an increase in threat frequency within the years 2015-2018. The maximum frequency occurs only for attacks that have an ascending threat frequency throughout the selected time span (Židková et al., 2020).

Generally defined categories of attacks we mentioned in the previous text are mapped to LTE threats. This mapping is also described in the Table 2, which shows us, that the greatest effort of the attackers is to process control, which can be a source of potential gain.

Detailed analysis of LTE related threats is provided in the following table Table 3.

Table 3. Comprehensive threat analysis vs severity vs attack type and vs attack category, source: authors

EBIOS categories by objectives	LTE attacks – a dreaded event	Degree of severity (negligible, restrictive, significant, critical)	Degree of probability (minimal, significant, strong, maximal)
Process control	Botnet	critical	Strong

	Exploit kit	critical	Strong
	Malware	critical	Significang
	Web Application Attack	negligible	Strong
Spying and stealing data	Cyber Espionage	negligible	Significant
Process disruption	Denial of Service	significant	Maximal
	Web Based Attack	negligible	Maximal

4.2. **Risk diagnostics**

The diagnosis of LTE risks (see next Figure) shows that critical threats are Botnet (R1), Exploit Kit (R2), Malware (R3), DoS (R6), and Web-Based Attack (R7). None of the critical threats is spying on and stealing data, but on the contrary, threats are targeted to process control (R1, R2, R3) and process disruption (R6, R7). LTE risk diagnosis is shown in the following Figure.
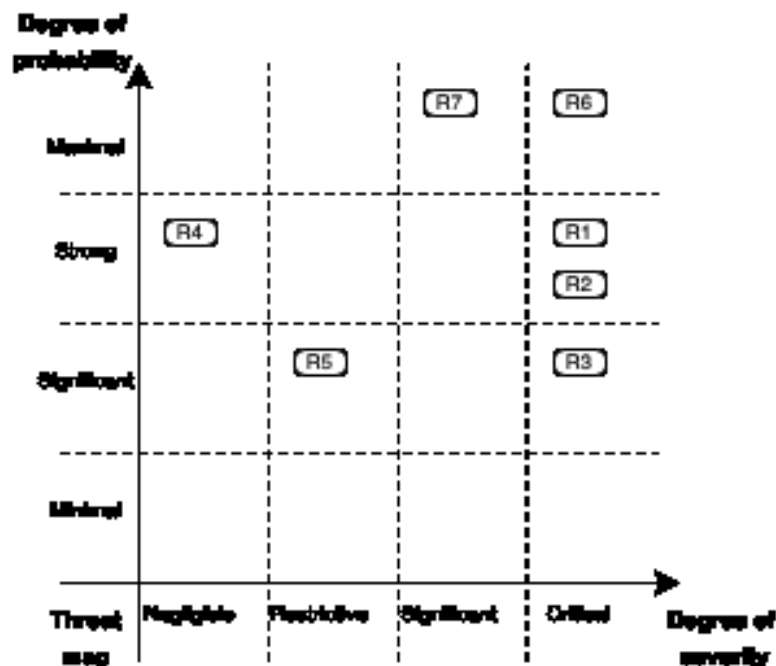
Figure 3. LTE risk diagnostics, source: authors

5. **Conclusions**

Security issues of IoT devices are increasing with their massive use in everyday life. This topic has been addressed by many researchers seeking to increase the privacy of the IoT, and the security of communication technologies is one of the key aspects. The excessive number of incoming IoT devices over the next years would cause serious security and privacy concerns if we continue utilizing the existing IoT device design and architecture.

The diagnosis of LTE risks provided in this paper showed us that critical threats are Botnet (R1), Exploit Kit (R2), Malware (R3), DoS (R6), and Web-Based Attack (R7). None of the critical threats is spying on and stealing data, but on the contrary, threats are targeted to process control (R1, R2, R3) and process disruption (R6, R7).

In context with the results of the security risk analysis, several actions can help to secure IoT devices:

- **Disable data sharing with unused services** - most systems allow users to enable or disable sharing of specific services, such as location sharing.

- **Use longer and more complicated passwords. Use a separate password for each device** - setting more complicated passwords prevents you from getting passwords with brute force, and almost makes it impossible to extract a password during the service operation. Reusing passwords is not a good idea. With Password Manager, you can track all your passwords.

- Never receive files or messages from untrusted devices - untrusted messages can contain an attack against the device.

- **Consider changing your Wi-Fi settings to not automatically connect** - this gives you more control over when and how your device uses Wi-Fi networks publicly

- **Do not stay logged into accounts permanently** - sign out when you're finished using your account

- Avoid public Wi-Fi networks. Or use a virtual private network (VPN) to access your online accounts regularly via Wi-Fi hotspots - you may want to manage your IoT device through a mobile device in a city café. If you use public Wi-Fi - which is generally not a good idea based on the analysis - use VPN.

- **Use two-factor authentication** - for example, a one-time code sent to your mobile phone - can keep attackers away from your device. If IoT device applications offer two-factor authentication, use it.

The IoT security is important topic for all IT users/researches, because almost everyone is faced to various IoT technologies. All losses related to the security breaches can affect not only private companies but public economic also.

6. Acknowledgement

7. References

Cichonski, J., Franklin, J. M., & Bartock, M. (2017). Guide to LTE security (NIST SP 800-187; p. NIST SP 800-187). National Institute of Standards and Technology. Retrieved January 07, 2020, from https://doi.org/10.6028/NIST.SP.800-187

EBIOS. (n.d.). Publication: EBIOS — Expression des Besoins et Identification des Objectifs de Sécurité. ANSSI. Retrieved January 08, 2020, from https://www.ssi.gouv.fr/guide/ebios-2010-expression-des-besoins-et-identification-des-objectifs-de-securite/

ENISA. (2016). ENISA Threat Landscape 2015. ENISA Publishing. Retrieved February 27, 2020, from https://www.sbs.ox.ac.uk/cybersecurity-capacity/system/files/ENISA%20Threat%20Landscape%202015.pdf

ENISA. (2017). ENISA Threat Landscape Report 2016. ENISA Publishing. Retrieved February 27, 2020, from https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2016

ENISA. (2018). ENISA Threat Landscape Report 2017. ENISA Publishing. Retrieved February 27, 2020, from https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2017

ENISA. (2019). ENISA Threat Landscape Report 2018. ENISA Publishing. Retrieved February 27, 2020, from https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018

Frankel, S., Eydt, B., Owens, L., & Scarfone, K. (2007). Establishing Wireless Robust Security Networks. National Institute of Standards and Technology, U.S. Department of Commerce. Retrieved March 06, 2020, from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-97.pdf

Hanes, D. (2017). IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things (1 edition, Vol. 2017). Cisco Press.

Hegerová, M., & Dostálek, Libor. (2016). Mobilní sítě a jejich bezpečnost. Ústav aplikované informatiky, Přírodovědecká fakulta Jihočeské univerzity v Českých Budějovicích. Retrieved February 10, 2020, from https://www.prf.jcu.cz/data/files/344/433/2481mobilnisite.pdf

Keenan, T. (2017, July 6). Securing the Internet of Things. Hiring Headquarters. Retrieved February 22, 2020, from https://www.upwork.com/hiring/data/securing-internet-things/

Kidston, D., Tang, H., Mason, P., & Li, L. (n.d.). Mitigating Security Threats in Tactical Networks. NATO OTAN. Retrieved March 16, 2020, from https://pdfs.semanticscholar.org/5a53/5491fa5e3bab4e484199ede8d9e6aee657da.pdf

Mohamad Noor, M., & Hassan, W. (2018). Wireless Networks: Developments, Threats and Countermeasures.

Padraig, S. (2016, November 29). Understanding IoT Security – Part 1 of 3: IoT Security Architecture on the Device and Communication Layers - IoT Analytics. Retrieved February 28, 2020, from https://iot-analytics.com/understanding-iot-security-part-1-iot-security-architecture/

Qiu, Y., Zhou, J., & Deng, R. (2006). Security Analysis and Improvement of Return Routability Protocol. In M. Burmester & A. Yasinsac (Eds.), Secure Mobile Ad-hoc Networks and Sensors (pp. 174–181). Springer. https://doi.org/10.1007/11801412_16

Sinha, S. R., & Park, Y. (2017). Building an Effective IoT Ecosystem for Your Business (1st ed. 2017 edition, Vol. 2017). Springer.

Židková, N., Maryška, M. & Doucek, P. (2020, April 02-03). Security of Wi-Fi as a Key Factor for IoT [Paper presentation]. Hradec Economics Days, Hradec Králové (pp. 906-914). https://doi.org/10.36689/uhk/hed/2020-01-101

VULNERABILITIES OF SCHOOL WEBSITES IN V4 COUNTRIES

Jaroslav Svoboda, Jiří Georgiev

Faculty of Informatics and Statistics
Prague University of Economics and Business
jaroslav.svoboda@vse.cz, jiri.georgiev@vse.cz

Keywords

security, CMS, HTTPS, vulnerabilities, schools

Abstract

Software vulnerabilities represent a significant threat to the ICT security of educational institutions. A successful attack can endanger private information of students and employees, and expose the institution's data. This contribution maps the distribution of most common HTTPS and CMS vulnerabilities affecting school webs across all members of Visegrád Group. 36 064 websites of educational institutions were scanned and analysed. The most common HTTPS vulnerabilities are BREACH, SWEET32, BEAST and potentially LUCKY13. Approximately 10 % of WordPress installations and more than half of Joomla and Drupal installations suffer from one or more vulnerabilities. File security.txt which could be used to inform the institutions about such vulnerabilities is provided only by a handful of institutions. Although relative usage of CMS varies among countries, vulnerability findings are very similar. Recommendation to mitigate the most common vulnerabilities is given.

1. Introduction

Software vulnerabilities are an inevitable phenomenon which enables successful hacker attacks. These attacks can result, for example, in an unauthorised modification of data, their theft, unwanted encryption by ransomware or other types of malware infections. When an educational institution is affected, it is a very unfortunate situation which complicates not only the administrative or teaching activities of the school but also might lead to significant financial losses. Thus, a successful attack may have an impact on:

- *Reputation:* in the eyes of students, parents, and other stakeholders

- *Economy:* cost of data recovery or loss, ex-post securing of systems

- *Operation:* preventing and hindering the institution's activities

- *Legal:* risk of prosecution by aggrieved stakeholders ('Cyber security and universities: managing the risk', 2013)

The outcome of a successful attack might even be more severe because primary users of these websites are pupils and parents. Both groups trust the institution and heavily rely on the provided information, that might make them more susceptible to various attacks implementing social engineering.

As ICT specialisation is usually not the primary goal for educational institutions, and headmasters have a wide range of responsibilities, ICT security and its requirement of periodical attention might not be satisfied. This results in suboptimal decisions, for example, preference of free solution instead of sustainable and secure ones (Svoboda and Georgiev, 2019).

In our previous research (Svoboda and Georgiev, 2020) about 40 000 domains were scanned, and quality of connection security has been assessed in terms of HTTPS usage, supported SSL/TLS protocols, key strength and used certificates. We found that poor implementation is widespread among Visegrád Group countries, and only 26.5 % of websites have HTTPS implemented in such a way that can be considered good enough by our definition.

The focus of this paper is on HTTPS vulnerabilities, used content management systems (CMS) and their vulnerabilities, and implementation of security.txt file on the web. Thus, the goal is to assess how HTTPS implementations are susceptible to most known vulnerabilities. We also inspected websites for the type of CMS used and compared it with a database of its known vulnerabilities. We compare those findings with the results from the pool of Alexa top 10 million websites (*Usage statistics of content management systems*, 2020).

The structure of this paper is as follows. Description of various vulnerabilities and security measures are described in the first section. Following section concerns methodology for data collection and analysis used. Results are presented in section 3, and the paper is concluded in section 4.

1.1. HTTPS vulnerabilities

HTTPS had over time many vulnerabilities which were either the fault of the protocol itself or its implementation. The former is usually fixed by addition or by an update of the protocol. The latter must be fixed by authors of the particular software, and these fixes are then released in a form security update. Both cases require action from administrators of web servers who must follow news in ICT security and update their software on a regular basis. In this paper, we examined these following vulnerabilities.

The Heartbleed Bug discovered in 2014 is a bug in OpenSSL library which affects both server and client instances. The bug lets clients receive longer Heartbeat Request messages than the payload actually is, therefore, the received message consists not only of the payload but also whatever content of the memory buffer comes after. Exploitation can lead to a leak of private data such as passwords, usernames, emails etc. Heartbleed was exploited by hackers in several cases. It can be mitigated by the installation of a patched version 1.0.1g or newer.

CCS Injection is another OpenSSL bug that affects processing of ChangeCipherSpec messages. It can be exploited through the use of man-in-the-middle (MITM) attack, which allows decrypting of the data in transit because the attacker can enforce the use of weak keying material. It is limited to the cases where both client and server use vulnerable versions of the library. It can be mitigated by updating to version 1.0.1h or newer.

Ticketbleed is similar in spirit and implications to the Heartbleed vulnerability. Ticketbleed allows the attacker to extract data from uninitialised memory. It affects the TLS stack of F5 products.

Return Of Bleichenbacher's Oracle Threat (ROBOT) attack is as the name suggests a return of vulnerability from 1998. It affects RSA encryption key exchange of many vendors such as F5, Citrix, Cisco, IBM, or Symantec. It is possible to mitigate the issue by switching RSA encryption off.

Authentication Gap vulnerability: SSL and TLS allow to renegotiate a communication channel with a new handshake, but this is not cryptographically bound to the original handshake. Therefore, the attacker can intercept clients SSL/TLS connection and insert his data. This vulnerability affects many

SSL 3.0 and TLS implementations from various vendors. It can be mitigated by updating to patched versions of the software.

Secure Client-Initiated Renegotiation vulnerability affects OpenSSL 0.9.8k and earlier, which allows insecure client-initiated renegotiation.

Compression Ratio Info-leak Made Easy (CRIME) attack described in 2012 affects data (HTTP requests) which are compressed prior to encryption when being sent over TLS and SPDY protocols. It is most useful in stealing authentication cookies and therefore, hijacking sessions. Mitigation can be achieved by turning off TLS compression or SPDY header compression.

Browser Reconnaissance and Exfiltration via Adaptive Compression of Hypertext (BREACH) is an attack described in 2013 based on CRIME affecting gzip and DEFLATE compression used by HTTP. The attack, which revealed emails, passwords, and other private data, was demonstrated to take only 30-60 seconds. Probably the best mitigation is disabling the HTTP compression.

Padding Oracle On Downgraded Legacy Encryption (POODLE) is an attack which takes advantage of fallback to SSL 3.0. The mitigation is straightforward - disabling SSL 3.0, which has been already deprecated or updating software to versions which implement TLS_FALLBACK_SCSV (Moeller and Langley, 2015). There is also a version of this attack against TLS which is successful only because some implementations fail to validate the padding as TLS standard mandates properly.

Sweet32 attack exploits TLS connections which use weak block cyphers such as DES or Triple DES. Its principle is based on the birthday attack. It does take approximately two days to cause a collision and extract session Cookie (Bhargavan and Leurent, 2016). DES and Triple DES have been already deprecated. Therefore if not used, it is not possible to carry out the attack.

Factoring RSA Export Keys (FREAK) is an exploit which affects SSL/TLS protocols because of limited 512 or less bit-long RSA_EXPORT keys. This was caused by limitations imposed by US cryptography regulations.

Decrypting RSA with Obsolete and Weakened eNcryption attack allows, thanks to a bug, exploiting support for obsolete SSLv2 protocol on servers otherwise running newer protocols but sharing the same public key with this deprecated protocol. The best mitigation is to disable SSLv2 on the affected server.

Logjam is a vulnerability of 512-1024-bit keys used for Diffie-Hellman (DH) key exchange. It can be used in MITM attack which forces downgrade to 512-bit DH. Mitigation can be achieved using the elliptic-curve Diffie-Hellman. However, the real-life impacts of vulnerability were disputed by (Ronen and Shamir, 2016).

Browser Exploit Against SSL/TLS (BEAST) exploits the vulnerability of cypher block chaining in TLS 1.0. At the time, the vulnerability was mitigated by the use of RC4 cypher, which has its own weaknesses. Because the attack only affects browsers, the mitigation has to be done by updating them.

Lucky Thirteen attack is a timing side-channel attack against the message authentication code (MAC) check stage of TLS described in 2013. It is one of the more theoretical attacks because of its delicate nature of differential timing of server response. The name refers to TLS header and sequence number length of 13 bytes. The vulnerability was mitigated in 1.0.1d, but it is advisable not to use CBC cyphers at all.

RC4 is a stream cypher used in SSL/TLS protocols which was prohibited by (Popov, 2015) and is not part of TLS 1.3 specification at all because of proven vulnerabilities by (Sepehrdad, Vaudenay and Vuagnoux, 2011) and (Isobe *et al.*, 2014).

1.2. CMS vulnerabilities

Many organisations are using Content Management Systems (CMS). These systems enable non-technical personnel to edit the website without any prior knowledge of web technologies. There are many CMS's such as WordPress, Joomla or Drupal, which are free software and therefore, are widely used. All mentioned CMS's allow the use of plugins which add additional functionality such as fancier photo gallery or SEO optimisation, and themes which allow customisation of visual aspects. One website can use many plugins; therefore, the attack surface is widened because those plugins are separately developed software with their own bugs and vulnerabilities. Therefore, such software needs to be regularly maintained and updated because attackers can take advantage of a vulnerability in a matter of hours. They can, for example, steal data from the website, add malicious scripts to the source code of the website, use infected web for malware distribution or for phishing. Report (Avillez, 2020) shows that 56 % of all CMS installations were outdated at the point of infection, specifically WordPress in 49 %, Drupal in 77 % and Joomla 90 % of cases. It also states that WordPress is the most targeted CMS. Hence running the current version of CMS is a significant factor in the security of the whole website. CSIRT.CZ, National Cyber Security Response Team of the Czech Republic, provides scanning of .cz domains running CMS for free (*Metodika skenování redakčních systémů na webových stránkách v doméně .CZ - CSIRT*, 2017) and each year scans only tens of websites.

1.3. Security.txt

Although still Internet-Draft (Shafranovich and Foudil, 2020), the security.txt file is a useful source of information for anybody who finds a vulnerability or a bug. Its purpose is to provide a way to report discovered security vulnerabilities to personnel responsible for the website's security. Standard similar to robots.txt in current version 9 covers eight directives of which Contact and Encryption are the minima necessary for secure reporting of discovered vulnerabilities. The standard for security.txt mandates use of HTTPS protocol. The file must be placed in a *.well-known* directory (Nottingham, 2019) which resides in the root of the website and for legacy reasons it also might be placed in the top-level path or redirected to the mentioned primary location. It is possible to generate personalised security.txt file on the website securitytxt.org. This contribution also maps the usage of this standard among educational institutions.

2. Method and data

Primary data sources of websites are official databases: Czech database Stistko, Hungarian database Oktasi, Polish database Wykaz and Slovakian database CVTI. In addition to these sources, web-scraping has been used for InspIS PORTÁL of Czech School Inspectorate and Institutional Register of the Public Education Information System in Hungary. Additional database from The Ministry of Education, Science, Research and Sport of the Slovak Republic (RIS - portal) was also merged. This was very important due to the poor quality of data in the official database provided by the Ministry of Education, Youth and Sports of the Czech Republic and the Hungarian Office of Education database file. Especially contact information and website addresses were either partially missing or were unavailable at all in data sources of all countries. In the case of Slovakia, the data were not up to date in the primary source.

Testssl.sh (Wetter, 2019) was used for SSL/TLS vulnerabilities detection. CMSeeK (r3dhax0r, 2020) was used to get the type and basic info about running CMS and its version. Vulnerabilities of the top three CMS's were sourced from the web CVE Details.

Results were sourced from a total of 8000 domains in Czechia, 2993 domains in Hungary, 21 466 domains in Poland and 3665 domains in Slovakia.

3. Results

Table 1 shows the distribution of the most common HTTPS vulnerabilities in websites with functional HTTPS. It is important to note that the majority of websites cannot be assessed for these vulnerabilities because they do not offer HTTPS at all (Svoboda and Georgiev, 2020). Therefore, no kind of attack is needed because all data are transferred in plaintext. The most common vulnerabilities are BREACH, SWEET32, BEAST and LUCKY13. 2-5 % of websites still offer the use of RC4 cypher even though it has been prohibited for more than five years at the time of writing.

	CZ, N=3324	HU, N=834	PL, N=4689	SK, N=567
Heartbleed	0.06 %	0.00 %	0.00 %	0.00 %
CCS	0.06 %	0.00 %	0.00 %	0.00 %
Ticketbleed	0.00 %	0.00 %	0.00 %	0.00 %
ROBOT	0.09 %	0.12 %	0.51 %	0.00 %
Secure Renegotiation	0.00 %	0.00 %	0.00 %	0.00 %
Secure Client Renegotiation	1.14 %	0.12 %	0.15 %	0.00 %
CRIME	0.06 %	0.00 %	0.06 %	0.00 %
BREACH	74.85 %	45.32 %	78.59 %	83.77 %
POODLE	0.00 %	0.00 %	0.00 %	0.00 %
TLS_FALLBACK_SCSV	4.51 %	0.24 %	0.92 %	1.59 %
SWEET32	41.19 %	39.69 %	27.47 %	56.61 %
FREAK	0.03 %	0.00 %	0.00 %	0.18 %
DROWN	0.00 %	0.00 %	0.00 %	0.00 %
LOGJAM	0.03 %	0.00 %	0.00 %	0.18 %
BEAST	71.63 %	72.30 %	48.67 %	91.01 %
LUCKY13	99.31 %	91.25 %	98.36 %	98.77 %
RC4	2.95 %	5.52 %	2.30 %	6.17 %

Table 1: Occurrence of vulnerabilities by country

Only websites which have good enough implementation of HTTPS were included, i.e. those presenting a valid certificate and supporting TLS 1.0 or newer. These websites do not show visible errors in the browser of a visitor.

Table 2 shows the share of CMS systems and percentage of instances which, based on their detected version, have at least one known vulnerability. Quite alarming is the fact that more than half of Joomla and Drupal installations are not updated and therefore are exploitable. The most used CMS

is WordPress with around 10% of vulnerable instances, but thanks to the frequent use of insecure plugins which can be exploited, the vulnerability rate will most likely be higher than that. Only around 5% of educational institutions opted to use a different, lesser-known CMS system.

	CZ, N=8000		HU, N=2933		PL, N=21466		SK, N=3665		W3Techs, N=10M
CMS use	24.71		46.34		42.47		16.13		43.40
security.txt	0.55		0.10		0.00		0.00		0.11*
	CZ, N=1977		HU, N=1387		PL, N=9117		SK, N=591		W3Techs, N=4.34M
	MS	PV	MS	PV	MS	PV	MS	PV	MS
WordPress	67.83	9.99	53.57	6.86	55.44	10.68	56.35	9.61	63.20
Joomla	21.45	54.95	31.43	64.22	35.77	63.76	32.15	65.26	4.20
Drupal	5.31	46.67	9.66	64.18	2.19	68.00	5.75	58.82	2.80
Other	4.81		4.97		5.64		4.40		29.80

Table 2: CMS market share (MS, in %), potentially vulnerable (PV, in %) instances and use of security.txt

*out of Alexa top 1M websites (Bode, 2019)

File security.txt is used only by a few institutions. The content is usually simple Contact directive with an email address and a telephone number with short comments such as: "# If you have found a security threat on this site, please contact us.". In all analysed cases, security.txt files were not placed on the web by the institutions themselves, because the contact refers to the company which created the website and probably have been taking care of it. Only websites created by Solidpixels also included Encryption and Signature directives. Among V4 countries, usage of security.txt is most prevalent in the Czech Republic with 0.55 % of scanned websites using it.

4. Conclusion

Websites of 36 064 educational institutions of Visegrád group countries were scanned for HTTPS and CMS vulnerabilities. A major part of analysed websites with working HTTPS had at least one unresolved vulnerability in HTTPS implementation, vulnerable CMS version or both.

There are no major differences among V4 countries in the case of HTTPS and CMS vulnerabilities. Institutions in all four countries have significant drawbacks in web security. These results are in line with previous studies of connection security and improve our understanding of the poor quality of ICT solutions used by and sometimes also implemented by educational institutions. Such institutions provide services for tens to low thousands of pupils and parents and therefore are comparable to small companies. Furthermore, insufficient ICT security in privately-owned companies is a comparably marginal issue because customers can move to the competition. In the case of educational institutions, this is often not as easily done. The vast majority of educational institutions are public institutions, yet they differ in approaches to IT security, mainly because of missing unifying policies from directing ministry. Therefore, the state of ICT security depends on the awareness of the management of a particular institution. This is definitely not an ideal state, because the management qualified to assess, review or implement own security policies. Moreover, it is not clear if it is necessary to burden school management with it.

To mitigate most common vulnerabilities found in this research, the administrators should turn off HTTP compression, forbid old cyphers (DES, TDES, RC4) and CBC cyphers, and offer only TLS 1.2, ideally newer. The CMS should be periodically updated. All three most used CMS's support check for new versions and some sort of automatic updating system.

Our study is limited by the quality of available public data and scanning techniques. The results are not as detailed as could be because of a wide range of all the possible combinations of the CMS software, versions and configurations. The future research should focus on other types of HTTPS vulnerabilities and attacks such as HEIST attack (Vanhoef and Goethem, 2016), or vulnerabilities of installed CMS plugins and themes.

5. Acknowledgement

6. References

Avillez, E. (2020) 2019 Website Threat Research Report. Sucuri. Available at: https://sucuri.net/wp-content/uploads/2020/01/20-sucuri-2019-hacked-report-1.pdf.

Bhargavan, K. and Leurent, G. (2016) 'On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN', in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16. the 2016 ACM SIGSAC Conference, Vienna, Austria: ACM Press, pp. 456–467. doi: 10.1145/2976749.2978423.

Bode, C. (2019) Search the security.txt records found in the Alexa Top 1 Million websites - matix.io. Available at: https://matix.io/security-txt-search/ (Accessed: 18 April 2020).

'Cyber security and universities: managing the risk' (2013), p. 24. Available at: https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2013/cyber-security-and-universities.pdf.

Isobe, T. et al. (2014) 'Full Plaintext Recovery Attack on Broadcast RC4', in Moriai, S. (ed.) Fast Software Encryption. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 179–202. doi: 10.1007/978-3-662-43933-3_10.

Metodika skenování redakčních systémů na webových stránkách v doméně .CZ - CSIRT (2017). Available at: https://csirt.cz/en/metodika-skenovani-redakcnich-systemu-na-webovych-strankach-v-do/ (Accessed: 2 May 2020).

Moeller, B. and Langley, A. (2015) TLS Fallback Signaling Cipher Suite Value (SCSV) for Preventing Protocol Downgrade Attacks. RFC7507. RFC Editor, p. RFC7507. doi: 10.17487/RFC7507.

Nottingham, M. (2019) Well-Known Uniform Resource Identifiers (URIs). RFC8615. RFC Editor, p. RFC8615. doi: 10.17487/RFC8615.

Popov, A. (2015) Prohibiting RC4 Cipher Suites. RFC7465. RFC Editor, p. RFC7465. doi: 10.17487/rfc7465.

r3dhax0r (2020) Tuhinshubhra/CMSeeK. Available at: https://github.com/Tuhinshubhra/CMSeeK (Accessed: 29 June 2020).

Ronen, E. and Shamir, A. (2016) 'Critical Review of Imperfect Forward Secrecy', in.

Sepehrdad, P., Vaudenay, S. and Vuagnoux, M. (2011) 'Discovery and Exploitation of New Biases in RC4', in Biryukov, A., Gong, G., and Stinson, D. R. (eds) Selected Areas in Cryptography. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 74–91. doi: 10.1007/978-3-642-19574-7_5.

Shafranovich, Y. and Foudil, E. (2020) A File Format to Aid in Security Vulnerability Disclosure. Available at: https://tools.ietf.org/html/draft-foudil-securitytxt-09 (Accessed: 6 April 2020).

Svoboda, J. and Georgiev, J. (2019) 'Usage and security of email as major means of online communication of pre-primary, primary and secondary Czech educational institutions', Efficiency and Responsibility in Education. Available at: https://erie.v2.czu.cz/en/r-13320-conference-proceedings.

Svoboda, J. and Georgiev, J. (2020) 'Security of connection to school websites: a comparison of Visegrád Group member countries', Efficiency and Responsibility in Education.

Usage statistics of content management systems (2020). Available at: https://w3techs.com/technologies/overview/content_management (Accessed: 18 April 2020).

Vanhoef, M. and Goethem, T. V. (2016) 'HEIST: HTTP Encrypted Information can be Stolen through TCP-windows', p. 27.

Wetter, D. (2019) testssl.sh. Available at: https://github.com/drwetter/testssl.sh (Accessed: 13 April 2019).

SUSTAINABILITY AND PERFORMANCE MANAGEMENT

SUSTAINABILITY MANAGEMENT IN CZECH AND SLOVAK COMPANIES

Petr Petera, Jaroslav Wagner

Department of Management Accounting
Faculty of Finance and Accounting
University of Economics, Prague
petr.petera@vse.cz, wagner@vse.cz

Renáta Pakšiová

Department of Accountancy and Auditing
Faculty of Economic Informatics
University of Economics in Bratislava
renata.paksiova@euba.sk

Keywords

Abstract

The aim of this paper is to contribute to the knowledge about the status quo of sustainability management in for-profit companies domiciled in the Czech Republic and Slovak Republic. The data were gathered via questionnaire survey and statistically compared between both countries. The results suggest that the approach to sustainability management of respondents from both countries is rather similar than different and there are numerous common patterns. Regarding sustainability reporting, IFRS adopters report more sustainability information than non-adopters in annual reports. Companies, which include sustainability into their strategy and implement sustainability-related plans and programs, structure and systems and measure their environmental and social performance, also disclose more sustainability-related information.

1. Introduction

Sustainability is a highly topical issue and the amount of research grows accordingly. Moreover, in the context of the development of digital economy and industry 4.0 (Doucek & Hološka, 2019) it seems that the importance of sustainability topics will remain stable or even grow.

Numerous articles (e.g. Habek, 2017; Habek, Bialy, & Livenskaya, 2019; Horvath et al., 2017; Kedzior, Cyganska, & Syrrakos, 2020; Pelikanova 2019; Petera, Wagner, Paksiova, & Krehnacova, 2019; Tetrevova, 2018) were dedicated to sustainability in the Czech Republic, Slovak Republic as well as other Central and Eastern European countries. Nevertheless, these articles were primarily dedicated to the sustainability reporting whereas in this paper the attention is primarily aimed at sustainability management.

The basic research question of this paper is: to which extent are sustainability issues embedded in the strategy of Czech and Slovak Companies? Is the attention paid to sustainability issues and its priorities similar or substantially different in both countries? The need of a cross-country research was mentioned for example by Buhovac and Groff (2012, p. 92). The research questions are answered through the analysis of data obtained via original survey-based questionnaire research conducted in years 2018 and 2019.

Comparison of sustainability management in the Czech Republic and Slovak Republic may be interesting because there are cultural factors, which suggests that these two countries will be similar as well as factors suggesting the opposite. Hofstede insights (2020) compares cultures of various countries according to six criteria: power distance, individualism, masculinity, uncertainty avoidance, long term orientation, indulgence. The culture in the Czech Republic is supposed to be hierarchical, individualist, masculine, with high preference for avoiding uncertainty, pragmatic (i.e. adapt traditions to new conditions), and not indulgent (i.e. with cynicism and pessimism). The culture of Slovak Republic is supposed to be hierarchical (more than in the Czech Republic), with no preference regarding individualism, masculine (more than in the Czech Republic), with no clear preference regarding risk avoidance, strongly pragmatic culture, not indulgent (approximately same as the Czech Republic).

The approach used in the research presented in this paper is based on literature emphasizing the need of integration of sustainability with strategy and with other management control systems (Epstein & Manzoni, 1998; Epstein & Roy, 2001; Epstein, Buhovac, & Yuthas, 2015; Ghosh, Herzig, & Mangena, 2019; Morioka & de Carvalho, 2016).

According to the numerous scholars, sustainability should be integrated into vision and strategy of companies and supported by implementation of the relevant tools. Specifically, Epstein and Roy (2001) provide a framework for translation of sustainability strategy into action. According to this framework, strategy should be accompanied by plans and programs, which help to achieve sustainability related goals and objectives. Moreover, appropriate structure and management systems (e.g. ISO 14001, EMS) have to be implemented to translate strategy into action. Our questionnaire is based on this framework and we therefore investigate whether sustainability strategy is accompanied by the appropriate plans and programs (both in the area of environmental management and in the area of social management) and appropriate structure and systems (again both in the area of environmental and social responsibility). Moreover, we enrich this model with the issue of performance measurement and investigate if respondents measure their environmental and social performance.

Although sustainability reporting is not the main topic of this paper, some important facts regarding reporting are discussed. An overview of the status quo of reporting in annual reports is provided and consequently is discussed relation between amount of reporting and IFRS (International Financial Reporting Standards) adoption as well as correlation between the amount of reporting and intensity of sustainability management.

2. Methodology and data collection

The approach to data collection used in this research is a questionnaire survey. The questionnaire was prepared on the basis of an extensive literature review, pilot-tested and on the basis of obtained feedback, the final adjustments were made. The data were collected from for-profit, medium and large companies from all industrial sectors. The same questionnaire was distributed in the Czech Republic and Slovak Republic and data were collected from February, 2018 till April, 2019. Lengthy process of obtaining data from companies both in the Czech Republic and Slovak Republic, was the reason for such long period of data gathering. In the Czech Republic, 1,000 companies meeting the selection criteria were selected from Albertina CZ Gold Edition database. Of these companies, 984

companies were contacted, and 106 usable responses were obtained (response rate 10.77%). In the Slovak Republic, 475 companies were contacted, and 49 usable responses were obtained (response rate 10.32%).

The descriptive characteristics of our respondents can be found in Table 1.

Table 1 – Descriptive statistics of respondents

Statistics	CZ			SK			Total		
	FTE	Assets (thousands EUR)	Turnover (thousands EUR)	FTE	Assets (thousands EUR)	Turnover (thousands EUR)	FTE	Assets (thousands EUR)	Turnover (thousands EUR)
Mean	943	177,544	208,246	1,684	439,141	393,628	1,180	260,243	266,851
Median	430	42,323	44,356	860	86,001	117,419	578	49,531	57,248
Standard deviation	1,508	581,955	527,159	2,323	1,469,323	1,113,999	1,834	958,493	764,039
Skewness	4.998	7.677	4.984	3.631	6.001	5.868	4.378	8.064	6.878
Kurtosis	32.383	67.606	30.070	16.786	38.414	37.224	24.410	73.861	57.743

On the basis of Table 1 it is possible to summarize that respondents from Slovak Republic are larger companies according to all three key criteria (full time employees – FTE, assets and turnover). On the other hand, it is obvious that even companies from the Czech Republic are large enough so that they may have developed sustainability management.

For data analysis, methods of descriptive statistics, Mann-Whitney test and correlations are used. The descriptive statistics is employed to provide key information on the status quo of sustainability management and reporting. Furthermore, Mann-Whitney test is used to test differences in the amount of reporting between IFRS adopters and IFRS non-adopters. Last but not least, correlations between the indices measuring various facets of sustainability management are calculated.

3. Results

This chapter presents key results of the research. First, the attention is paid to the inclusion of sustainability into strategy and management. Second, a question of organizational placement of sustainability management is shortly addressed. Third, the issues of sustainability reporting are discussed.

3.1. Sustainability strategy and management

First, the issue of inclusion of sustainability topics into strategy of companies was investigated. This inclusion was measured through a group of eight questions and respondents were asked to which extent they address the areas of: (S1) growth of shareholder value, (S2) environmental issues, (S3) labor practices and decent work, (S4) human rights, (S5) local communities, (S6) anti-corruption, (S7) CSR-oriented suppliers' assessment, (S8) product responsibility. These questions were answered on a scale from 1 to 5, where 1 = only according to legislation, 2 = perceived as a strategic topic, but neither discussed nor formulated, 3 = formulated as part of the strategy, but not discussed across the company, 4 = formulated as part of the strategy and discussed across the company, 5 = it is the key strategic priority across company.

The results for the Czech Republic can be found in Table 2 and for the Slovak Republic in Table 3.

Table 2 – Inclusion of sustainability topics in strategy in the Czech Republic

Statistics	S1	S2	S3	S4	S5	S6	S7	S8
Mean	3.67	3.56	3.99	3.32	2.55	3.70	3.06	4.35
Median	4.00	4.00	4.00	4.00	3.00	4.00	3.00	5.00
Standard deviation	1.144	1.367	1.028	1.522	1.367	1.415	1.406	.947
Skewness	-0.409	-0.794	-1.160	-0.379	0.230	-0.762	-0.207	-1.580
Kurtosis	-0.836	-0.605	1.118	-1.321	-1.259	-0.727	-1.321	2.204

Table 3 – Inclusion of sustainability topics in strategy in the Slovak Republic

Statistics	S1	S2	S3	S4	S5	S6	S7	S8
Mean	3.33	3.51	3.86	3.08	2.90	3.71	3.22	4.04
Median	3.00	4.00	4.00	3.00	3.00	4.00	3.00	5.00
Standard deviation	1.449	1.356	1.099	1.525	1.358	1.472	1.177	1.338
Skewness	-0.174	-0.702	-1.079	-0.217	-0.172	-0.911	-0.377	-1.329
Kurtosis	-1.393	-0.652	0.913	-1.421	-1.256	-0.602	-0.663	0.525

It can be summarized that despite the fact that respondents from Slovakia are larger (see the descriptive statistics in Table 1), the inclusion of sustainability topics into strategy is more similar than different in the two investigated countries. Relatively surprising is the fact that the highest mean (and also median) value has product responsibility. Growth of shareholder value is not the top priority neither in the Czech Republic nor in the Slovak Republic. Moreover, both mean and median of this sustainability topic are lower in the Slovak Republic. Our results suggest that in the area of inclusion of sustainability topics into the strategy of companies, there are some differences, but these are not of critical importance.

The next part of our questionnaire investigated to which extent are implemented (1) formalized plans and programs, (2) structure and systems, (3) measurement in two critical areas of corporate social responsibility – environmental and social.

In the field of environmental management, we were asking about seven areas: (1) utilized materials, recycled materials and energy consumption, (2) impact of products on environment, (3) water withdrawals, (4) impact on biodiversity, (5) emissions, effluents and waste, (6) impacts from product distribution and employee travel, (7) environmental responsibility of suppliers. We asked our respondents about these seven areas in all three investigated management tools (plans and programs; structure and systems; measurement). The questions regarding plans and programs were answered on a scale from 1 to 6, where 1 = no plans and programs, 2 = plans and programs are prepared rarely, 3 = plans and programs are prepared repeatedly, 4 = plans and programs are prepared systematically for particular operations, 5 = plans and programs are prepared systematically for all operations, 6 = irrelevant area for our company. The questions regarding structure and systems were answered on a scale from 1 to 6, where 1 = no formalized system, 2 = no formalized system, implementation considered, 3 = fragments of a system exist, 4 = incomplete system in run, 5 = comprehensive system for company in run, 6 = irrelevant area for our company. Finally, questions regarding measurement were answered on a scale from 1 to 5, where 1 = no KPI, 2 = KPI are measured, 3 = KPI are measured and compared with planned or past values, 4 = KPI are measured and compared also with competitors, 5 = irrelevant area for our company.

In the field of management of social responsibility, we were asking about 14 areas: (1) health and safety management, (2) training of employees, (3) diversity of employees, (4) equal opportunities of employees, (5) discrimination, (6) impacts on society, (7) corruption, (8) safety of products, (9) impact of products on health, (10) energy-efficiency of products, (11) non-compliance of products, (12) supplier assessment for labor practices, (13) cooperation with local suppliers, (14) societal impacts of suppliers. Equally as in the case of environmental management, respondents were asked about these 14 areas in all three investigated management tools (plans and programs; structure and systems; measurement) and measurement scales were also same as in the case of environmental management.

It is above the scope of this paper to describe and analyze the results of all these questions in detail, but it is possible to mention that similar patterns were discovered between Czech and Slovak companies. For example, there are significant differences between the extent of measurement of individual indicators. While "impacts from product distribution and employee travel" is not measured by 67.4% of Czech respondents and 61.4% of Slovak respondents, indicator "emissions, effluents and waste" is not measured only by 17.2% of Czech respondents and by 12.5% of Slovak respondents. Similar patterns are exposed in the areas of plans and programs as well as structure and systems.

In order to provide an overall characterization of the approach of companies to the sustainability management, we computed indices as arithmetic means of the answers of respondents and results can be found in Table 4 and Table 5.

Table 4 – Description of indices, the Czech Republic

Index	Sustainabi-lity strategy I1	Environmental management			Social management		
		Plans&Pro-grams I2	Structure&Sys-tems I3	Measure-ment I4	Plans&Pro-grams I5	Structure&Sys-tems I6	Measuremen t I7
Mean	3.52	3.47	3.70	2.14	3.77	3.90	2.29
Median	3.63	3.59	3.83	2.18	3.96	4.08	2.14
Standard deviation	0.91	1.08	1.15	0.80	0.95	0.91	0.80
Skewnes s	-0.374	-0.383	-0.699	0.189	-0.459	-0.639	0.258
Kurtosis	-0.539	-0.830	-0.400	-0.795	-0.895	-0.503	-0.835

Table 5 – Description of indices, the Slovak Republic

Index	Sustaina-bility strategy I1	Environmental management			Social management		
		Plans&Pro-grams I2	Structure&Sys-tems I3	Measuremen t I4	Plans&Pro-grams I5	Structure&Sys-tems I6	Measuremen t I7
Mean	3.46	3.36	3.52	2.12	3.90	3.99	2.24
Median	3.50	3.33	3.64	2.29	4.21	4.29	2.29
Standard deviation	0.96	1.07	1.21	0.72	1.00	0.99	0.82
Skewnes s	-0.394	-0.416	-0.714	-0.135	-0.872	-0.974	0.178
Kurtosis	-0.484	-0.554	-0.528	-1.287	-0.161	0.029	-1.024

It is possible to sum up that in case of management and measurement of environmental issues, the mean value of indices is higher among Czech respondents. In case of management of social issues, indicators of plans and programs and structure and systems are higher for Slovak respondents, while indicator of measurement is slightly higher among Czech respondents. Nevertheless, we again see similar patterns in both countries. It seems that social issues are getting slightly more attention than environmental issues and index of structure and systems is in both countries slightly higher than index of plans and programs.

The mentioned finding regarding plans and programs and structure and systems is surprising because the relation between these two groups of management tools can be seen so that first there are plans and programs and consequently are developed structures and systems to support these plans. Further research is needed in this area because this study does not provide data necessary for answering such issues.

Finally, we computed correlations between indices and results can be found in Table 6 and Table 7.

Table 6 – Pearson correlations between variables, Czech sample

Index	I1	I2	I3	I4	I5	I6	I7
I1	1	0.696**	0.663**	0.583**	0.711**	0.621**	0.506**
I2		1	0.765**	0.710**	0.670**	0.547**	0.498**
I3			1	0.702**	0.583**	0.520**	0.509**
I4				1	0.610**	0.554**	0.628**
I5					1	0.894**	0.612**
I6						1	0.664**
** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).							

Table 7 – Pearson correlations between variables, Slovak sample

Index	I1	I2	I3	I4	I5	I6	I7
I1	1	0.760**	0.699**	0.464**	0.751**	0.700**	0.562**
I2		1	0.841**	0.580**	0.672**	0.648**	0.580**
I3			1	0.661**	0.680**	0.701**	0.614**
I4				1	0.358*	0.300*	0.646**
I5					1	0.906**	0.607**
I6						1	0.615**
** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).							

The results again support the fact that the approach in both countries is rather similar than different. Index of inclusion of sustainability into strategy (I1) is positively correlated with all other indices and all correlations are statistically significant at the 0.01 level. This indicates that companies, which incorporate sustainability issues into their strategy also develop plans and programs and structure and systems to support this strategy. A possible interpretation is that inclusion of sustainability-related topics into strategy is accompanied with their operationalization, i.e. companies go beyond formalistic proclamation and actively implement sustainability-related management.

The similarities between Czech and Slovak companies indicate that it would be possible to approach the data from both countries as one dataset and apply more sophisticated statistical methods to

investigate influence of sustainability strategy, plans and programs, and structure and systems on performance of companies in economic, environmental, and social area.

3.2. **Organization of sustainability management**

The organizational solution of sustainability management is an important issue and it is often suggested that establishment of separate corporate social responsibility department may be useful and strongly depends on size of a company. Our respondents were therefore asked if their company has a separate corporate social responsibility department and the results are discussed below.

This study found that 91 (85.8%) of Czech companies does not have a separate corporate social responsibility department and 15 companies (14.2%) does have such specialized department. In case of Slovak companies, the higher share of them with specialized CSR department was found – 11 companies (22.4%), but majority – 38 (77.6%) does not have specialized CSR department. We hypothesize that higher share of companies with a specialized CSR department among Slovak respondents may stem especially from two facts.

First, Slovak companies, which filled our questionnaire, are in average larger than Czech companies and the rationality of creating specialized CSR department grows with size of a company.

Second, large Slovak companies are usually influenced by their foreign parent companies and this may further increase the percentage of companies with specialized SCR department.

3.3. **Sustainability disclosure in annual reports**

Companies may report sustainability information in a large variety of media, e.g. standalone corporate social responsibility reports, web pages etc. Nevertheless, sustainability disclosure in annual reports is considered especially important. Annual reports are audited, and investors therefore regard them to be more trustworthy than other media.

The respondents were asked the following question: "Does your company issue information on environmental and social issues beyond the minimal legal requirements within annual report?".

This question was answered using the following scale: 1 = no, 2 = no, but we assume to do so from the next year, 3 = yes, the amount of information is up to one page, 4 = yes, the amount of information is over one and up to five pages, 5 = yes, the amount of information is over five pages. The answers were provided separately for environmental disclosure and for disclosure on social responsibility.

The frequencies of answers for the Czech Republic can be found in Table 8 and for the Slovak Republic in Table 9.

Table 8 – Reporting of sustainability information in the Czech Republic

Answer	Annual report – environmental information		Annual report – social information	
	Frequency	Percent	Frequency	Percent
1	53	50.0	58	54.7
2	3	2.8	2	1.9
3	25	23.6	28	26.4
4	14	13.2	11	10.4
5	11	10.4	7	6.6
Total	106	100.00	106	100.00

Table 9 – Reporting of sustainability information in the Slovak Republic

Answer	Annual report – environmental information		Annual report – social information	
	Frequency	Percent	Frequency	Percent
1	18	36.7	16	32.7
2	3	6.1	4	8.2
3	13	26.5	15	30.6
4	7	14.3	6	12.2
5	8	16.3	8	16.3
Total	49	100.00	49	100.00

In the rest of this subchapter two issues are analyzed in a more detail. First, the amount of sustainability reporting of IFRS adopters and non-adopters is compared. Second, correlation between the intensity of sustainability management (measured by indices I1–I7, see Table 4 and Table 5) and the amount of sustainability disclosure in annual reports is investigated.

3.3.1. Sustainability reporting of IFRS adopters and non-adopters in annual reports

Relationship between sustainability (corporate social responsibility) disclosure and IAS/IFRS is not intensively discussed in literature, nevertheless there are several studies dealing with this topic (e.g. Amelio, 2016; Lee, 2019; Negash, 2012). In this paper we investigate whether IFRS adopters report more sustainability-related information than non-adopters. In this case was not differentiated between Czech and Slovak companies and the results of analysis (Mann-Whitney test) can be found in Table 10.

Table 10 – Between-group comparison (IFRS non adopters and adopters), Czech Republic and Slovak Republic

Area of reporting	Group	Number of companies	Mean rank	Sig. (Mann-Whitney)
Environmental	non-IFRS	131	74.44	0.014
	IFRS	24	97.46	
Social	non-IFRS	131	73.02	0.001
	IFRS	24	105.17	

The results in Table 10 show that IFRS adopters disclose more sustainability-related information in their annual reports than IFRS non-adopters and the difference is statistically significant.

3.3.2. Correlation between intensity of sustainability management and the amount of sustainability-related disclosure in annual reports

Finally, we investigated correlations between the amount of reporting of environmental and social issues and indices defined in Table 4 and Table 5 (which measure the intensity of sustainability management). The detailed results are not presented here, but it is possible to summarize that all indices are statistically significantly and positively correlated with the amount of sustainability disclosure in annual reports. It is important to mention that according to some theories (Boiral, 2013), sustainability reporting may serve as a simulacrum to camouflage problems in sustainability management (i.e. companies with insufficient sustainability management should according to predictions of these theories report more). Our results support a different view – the amount of sustainability reporting grows with inclusion of sustainability into strategy and with implementation of plans and programs, structure and systems and measurement of environmental and social performance.

4. Conclusions

This paper provides a comparison of sustainability management in Czech and Slovak companies. The analysis is based on original questionnaire survey empirical research conducted among companies domiciled in the Czech Republic and Slovak Republic in years 2018 and 2019.

The results indicate that the approach of companies to sustainability management is more similar than different and it would be possible to analyze the obtained data also as one dataset to employ more sophisticated statistical methods for investigation of the influence of sustainability strategy, plans and programs, and structure and systems and measurement on performance in economic, environmental, and social area.

Regarding inclusion of sustainability-related topics into strategy, it was found that both in the Czech sample and Slovak sample the highest mean (and also median) value has product responsibility. Growth of shareholder value is not the top priority neither in the Czech Republic nor in the Slovak Republic (median in Czech sample is 4.0 and median in Slovak sample 3.0).

In order to make an overall characterization of the approach of companies to the sustainability management, we designed indices in several areas. Specifically, index of inclusion of sustainability issues into strategy, index of utilizations of sustainability-related plans and programs in area of environmental management, index of utilizations of sustainability-related structure and systems in area of environmental management, and index of sustainability-related management in area of environmental management. The same indexes were developed for area of management of social responsibility. We found that companies in the Czech and Slovak Republic include sustainability-related topics into strategy (see Table 2) and develop plans, programs, structure and systems to support this strategy (see Table 4 and Table 5). On the other hand, the results suggest that implementation is often only partial and substantial improvements are needed.

Correlation analysis proved that inclusion of sustainability-related topics into strategy is positively correlated with their operationalization through plans, programs, structure and systems., i.e. companies go beyond formalistic proclamation and actively implement sustainability-related management.

Interesting results were obtained in area of sustainability reporting. First, companies implementing IFRS in their annual reports disclose more sustainability-related information than IFRS non-adopters. Second, the amount of sustainability reporting grows with inclusion of sustainability into strategy and with implementation of plans and programs and structure and systems.

5. Acknowledgement

6. References

Amelio, S. (2016). The Connection between IAS/IFRS and Social Responsibility. Management Dynamics in the Knowledge Economy, 4(1), 7–30.

Boiral, O. (2013). Sustainability reports as simulacra? A counter-account of A and A plus GRI reports. Accounting Auditing & Accountability Journal, 26(7), 1036–1071.

Buhovac, A. R., & Groff, M. Z. (2012). Contemporary performance measurement systems in Central and Eastern Europe: A synthesis of the empirical literature. Journal for East European Management Studies, 17(1), 68–103.

Doucek, P., & Hološka, J. (2019). Digital economy and industry 4.0. Proceedings of the International Conference IDIMT-2019 Innovation and Transformation in a Digital World. Available at: https://idimt.org/wp-content/uploads/2019/08/IDIMT-2019-proceedings.pdf

Epstein, M., & Manzoni, J. F. (1998). Implementing corporate strategy: From tableaux de bord to balanced scorecards. European Management Journal, 16(2), 190–203.

Epstein, M. J., & Roy, M.-J. (2001). Sustainability in action: Identifying and measuring the key performance drivers. Long Range Planning, 34(5), 585–604.

Epstein, M. J., Buhovac, A. R., & Yuthas, K. (2015). Managing Social, Environmental and Financial Performance Simultaneously. Long Range Planning, 48(1), 35–45.

Ghosh, B., Herzig, C., & Mangena, M. (2019). Controlling for sustainability strategies: findings from research and directions for the future. Journal of Management Control, 30(1), 5–24.

Habek, P. (2017). CSR Reporting Practices in Visegrad Group Countries and the Quality of Disclosure. Sustainability, 9(12).

Habek, P., Bialy, W., & Livenskaya, G. (2019). Stakeholder engagement in corporate social responsibility reporting. The case of mining companies. Acta Montanistica Slovaca, 24(1), 25–34.

Horvath, P., Putter, J. M., Dagiliene, L., Dimante, D., Haldma, T., Kochalski, C., . . . Wagner, J. (2017). Status Quo and Future Development of Sustainability Reporting in Central and Eastern Europe. Journal for East European Management Studies, 22(2), 221–243.

Hofstede insights. (2020, May 1). Country comparison tool. Available at: https://www.hofstede-insights.com/country-comparison/czech-republic,slovakia/

Kedzior, M., Cyganska, M., & Syrrakos, D. (2020). Determinants of Voluntary International Financial Reporting Standards Adoption in Poland. Inzinerine Ekonomika-Engineering Economics, 31(2), 155–168.

Lee, W. J. (2019). Toward sustainable accounting information: Evidence from IFRS adoption in Korea. Sustainability, 11(4).

Morioka, S. N., & de Carvalho, M. M. (2016). A systematic literature review towards a conceptual framework for integrating sustainability performance into business. Journal of Cleaner Production, 136, 134–146.

Negash, M. (2012). IFRS and environmental accounting. Management Research Review, 35(7), 577–601.

Pelikanova, R. M. (2019). Corporate Social Responsibility Information in Annual Reports in the EU-A Czech Case Study. Sustainability, 11(1).

Petera, P., Wagner, J., Paksiova, R., & Krehnacova, A. (2019). Sustainability Information in Annual Reports of Companies Domiciled in the Czech Republic and the Slovak Republic. Inzinerine Ekonomika-Engineering Economics, 30(4), 483–495.

Tetrevova, L. (2018). Communicating CSR in High Profile Industries: Case Study of Czech Chemical Industry. Inzinerine Ekonomika-Engineering Economics, 29(4), 478–487.

PROPOSAL OF THE CONTINUOUS MODEL FOR INTERNAL AUDIT ASSESSMENT BASED ON COBIT 2019

Vlasta Svatá

Department of System Analysis
University of Economics, Prague
svata@vse.cz

Keywords

Abstract

In the light of digital society internal audit (IA) represents one of the three lines of defence aiming to decrease enterprise risks and improve controls. Establishing IA departments is a core condition for different types of assurance, but in the same time IA department itself should be subject of quality assessment. The article discusses different models for IA assessment and introduces the proposal for application of continuous model of assessment based on COBIT 2019.

1. Introduction

Internal auditing is an independent, objective assurance and consulting activity designed to add value and improve an organization's operations. It helps an organization accomplish its objectives by bringing a systematic, disciplined approach to evaluate and improve the effectiveness of risk management, control, and governance processes (IIAa, 2019). The main aim of internal auditing is thus to provide a written assessment to the board of the effectiveness of the organization's system of internal control and risk management. Many companies have made large investments in internal auditing in recent years and they want to ensure that the internal audit function is adding value to their organizations. Therefore, there is a need to assess the performance of the internal audit function itself. In the same time there exist the whole set of different approaches how to provide this assessment without clear explanation of what approach is the best in relation to specific context of the organization. Therefore this article, which is the output of qualitative research, attempts to answer two main questions:

- What are the main models enabling the internal audit (IA) quality performance assessment and how they differ each other?

- How COBIT 2019 supports the application of the IA quality assessment based on continuous model?

The core components of qualitative research were personal experience, analysis of both the obligatory and voluntary standards and guidelines and knowledge received through the case studies of COBIT 2019.

2. The Regulatory Background for IA Assessment

This need of the assessment of the IA function within the enterprises is mainly regulated by The IIA's International Standards for the Professional Practice of Internal Auditing (Standards). These Standards are published in the document International Professional Practices Framework (the IPPF). In accordance with the IIA's Standards, "The chief audit executive must develop and maintain a quality assurance and improvement program that covers all aspects of the internal audit activity" (Standard 1300). In order to provide more detailed guidance in establishing and maintaining a quality assurance and improvement program Internal Audit Foundation published Quality Assessment Manual for the Internal Audit Activity in 2017. This document deals with the need of the conformance to all mandatory elements of the IPPF and their demonstration to the stakeholders which are in responsibility of Chief audit executives (CAEs). The only way to meet these expectations is with „a comprehensive quality assurance and improvement program (QAIP)". This program has two dimensions (IIAb, 2017):

1. Dimension of the organizational aspects of the QAIP assessment that includes different levels of assessment that are identified by the help of assessment stakeholders and periodicity. This dimension has in documents of IIA priority and is focused on three main type of assessment that are compliant with the IA Standards:

    a. **Ongoing monitoring of performance** - is a core part of internal assessment. The main aim of it is to address conformance with the Standards: 2200 Engagement Planning, 2300 Performing the Engagement, 2400 Communicating Results and 2500 Monitoring Progress. These Standards are intended to address quality on an audit-by-audit basis and relate primarily to separate engagement activities.

    b. **Periodic internal assessments** – comparing previous type of assessment periodic internal assessment address conformance with Standards as well, but it provides more holistic, comprehensive review of the Standards. It is important in the early stages of establishing QAIP and according to successful IA practice should be repeated at least annually. Examples of Standards important for this type of assessment are 1000 Purpose, Authority and Responsibility, 1100 Independence and Objectivity, 1200 Proficiency and Due Professional Care, 2000 Managing the Internal Audit Activity and Code of Ethics. The way of assessment can be self-assessment or combination of self assessment and external assessment.

    c. **External assessments** – the aim is very similar to the periodic internal assessment but its additional inherent part is the feedback from stakeholders (board, audit committee and management of the organization). The communication with them should be conducted by a qualified, independent assessor or assessment team from outside the organization. It should be provided at least once every five years. Thus the impact is on external assessment and communication with stakeholders.

2. Dimension of the assessment scales (methods of assessment) is less discussed as the scales (methods) are not to a large extend defined in Standards. They are dependent on the context of the assessment. This context takes in account the strategy and quality of both the subject of the engagement and assessor. For instance internal auditing may be less mature in emerging countries, privately held (not listed) companies, not-for-profit organizations, small companies, and organizations with a relatively new internal audit activity. At the same time, many mature internal audit activities that are generally in conformance with the Standard*s* and the Code of Ethics look for ways to provide context to the operation of their activity.

3. Models for IA Quality Performance Assessment

There exist a lot of different documents dealing with the problem of how to assess the quality of internal audit function and thus to be compliant to Standard 1300. Examples of the relevant documents are (IIAb, 2017), (IPPF, 2013), (IIAc, 2020), (EU, 2013), (Schneider, 2010), (ISACA, 2018). After analysis of these documents I came to the conclusion that basically the different approaches to IA assessment fall into three categories of models:

- Traditional models
- Maturity models
- Capability models

3.1. Traditional Models

Traditional models are very important for ongoing monitoring. As noted in Implementation Guide 1311 – Internal Assessments, ongoing monitoring mechanisms may include:

- Checklists or automation tools to provide assurance on internal auditors' compliance with established practices and procedures, mainly 2200 etc.

- Feedback from internal audit clients and other stakeholders regarding the efficiency and effectiveness of the internal audit team

- Staff and engagement key performance indicators (KPIs), such as the number of certified internal auditors (CIAs) on staff, their years of experience in internal auditing, the number of continuing professional development hours They earned during the year, timeliness of engagements, and stakeholder satisfaction.

- Other measurements (project budgets, timekeeping systems, and audit plan completion)

Example of the checklist representing this traditional model is available in Table .

3.2. Maturity Models

Maturity models are recommended to be used within the periodic internal assessment as the relevant Standards representing the criteria for assessment do not prescribe an assessment scale. Furthermore it is expected, that the assessor will take in account the context of the organization. Therefore the checklists that are typical for traditional approach should be accompanied by more complex model. The results of the assessment based on maturity model should be confirmed by internal self-assessment process and eventually assessed by an external qualified, independent assessor or assessment team.

The idea of maturity models is part of Capability Maturity Model Integrations (CMMI) which is the product of CMMI Institute. The main aim of the model is to enable organizations to assess and improve their processes and activities. CMMI offers two representations/models for process improvement. The first one is staged representation which is the model of discrete process improvement. It means that multiple processes are checked at once. Separate levels are called maturity levels. The second representation is continuous model. It is a model of continuous improvement of individual processes and it deals with capability levels (see next chapter).

Table 1: Checklist for ongoing monitoring of internal assessment (Pempal, 2016, p 25)

Standard	Checklist
2200 Engagement Planning	a. Are individual internal audit engagements adequately resourced and properly monitored? b. Are the internal auditors familiar with the processes under review? c. Have critical risks been identified during the preliminary survey? d. Will the audit objectives allow auditors to provide assurance? e. Is the scope sufficient to satisfy the audit objectives? f. Will the audit program allow internal auditors to achieve the audit objectives and reach a conclusion? g. Have auditees been informed about the planned audit? h. Were the objectives clearly explained to auditees during the kick-off meeting?
2300 Performing the Engagement	a. Are all executed steps properly documented? b. Is the prescribed methodology being applied and are appropriate audit techniques being used? c. Have the internal auditors properly assessed auditees' procedures with regard to the processes under review? d. In the absence of auditees' procedures, have the internal auditors discussed with the auditees the assessment criteria that should be used? e. Is the obtained evidence sufficient to express an opinion? f. Do internal auditors differentiate between critical and less critical findings? g. Were findings immediately communicated and discussed with the auditees? h. Has the work program been carried out as intended? i. Are changes to audit objectives, scope and work program justified and properly approved?
2400 Communicating Results	a. Were the findings and final conclusion presented to the auditees at a closing meeting? b. Do the recommendations address the root cause of the findings? c. Are the recommendations practical? d. Does the audit achieve its objectives of being able to issue negative or positive assurance? e. Has a draft report been sent to auditees, allowing them to review and comment on the findings and recommendations? f. Have the internal auditors incorporated the auditees' comments? g. Do internal auditors agree on the action plan? h. Is the audit report accurate, objective, clear, concise, constructive and timely? i. Has the audit report been signed according to the relevant policies? j. Have audit objectives been achieved within allocated resource budgets and by agreed target dates, as much as possible?
2500 Monitoring Progress	a. Have the internal auditors monitored whether the deadlines of the action plan were respected? b. Have the internal auditors assessed whether a follow-up audit may be needed? c. Have follow-up activities been duly executed by the internal auditors?

Each maturity model must have three dimensions:

- Identification of the processes (activities, components) that are subject of assessment

- Identification of the maturity levels

- Identification of the indicators (key characteristics) that can be used for maturity levels assignment

There exist many different examples of maturity models that can be applied on to IA quality assessment. They are tailored to the needs of this type of assessment and furthermore can be adapted by an organization to provide additional insight into maturity levels for specific internal audit processes or elements of infrastructure. Two examples of such adjustments are shown in the Table as evidence of the possibility of adapting the original maturity model. The adjustments relate to the three key dimensions of the model mentioned above.

Table 2: Examples of maturity model modification for the purpose of IA assessment

Maturity model of IPPF (2013)	Maturity model of IIAd (2017)

IA processes	Services & Role of IA People Management Professional Practices Performance Management & Accountability Organizational Relationships & Culture Governance Structures	Maturity Recruiting, On-Boarding, and Staff Development Risk Assessment and Annual Audit Planning Execution of Internal Audit Methodology Use of Information Technology Reporting and Monitoring
Maturity levels	Level 5 – Optimizing Level 4 – Managed Level 3 – Integrated Level 2 – Infrastructure Level 1 - Initial	Level 5 - Optimized Level 4 - Managed Level 3 - Defined Level 2 - Repeatable Level 1 - Initial
Indicators	Strategic Planning/Financial Management Customer/Stakeholder Expectations Risk Metrics Human Capital Process Management and Self assessment	Policy Methodology and process People Systems and informatik Communication and reporting

3.3. Capability Models

Capability models in general are important part of management model called Enterprise Governance which was initiated by the set of bankruptcy scandals at the turn of the millennium (e.g. Enron). In short its main features are

- impact on stakeholders needs, mainly benefit realization, risk management and optimization of the risk

- performance management

- importance of information technology (IT)

The combination of these three fundamental organizational management goals was the basis for the emergence of the EGIT – Enterprise Governance of IT. External assessment of IA performance seems to be good object for capability model application. The reasons are next:

- being part of Enterprise Governance it takes in account stakeholders needs

- performance management is an essential part of EGIT

- the importance of IT within IA performance - even to the fact that it is not directly addressed in Standards -   still grows

One of the biggest promoters of EGIT is ISACA – Information Systems Audit and Control Association. In order to support and enable organizations to implement and operate EGIT it developed and published framework COBIT – Controls Objectives for IT. Currently the sixth version of COBIT is available named COBIT 2019.

4. Proposal for Application of Continuous Model for IA Assessment

Compared to the previous version COBIT 5 COBIT 2019 comes with a new approach to performance management called COBIT performance management (CPM). It is a general term for all concepts and methods (such as capability levels and maturity levels) that are influenced by the current development in this area, especially CMMI 2.0 methodology. It expresses how well the governance and management system and all the components of an enterprise work, and how they can be improved to achieve the required level. It is an integral part of the COBIT framework (ISACA, 2019).

One of the most important part of COBIT 2019 is COBIT Core Model. It identifies 40 governance and management objectives that can contribute to the achievement of enterprise goals. One of this objective is MEA04 Managed Assurance ("MEA" identifies the domain of the objective which is Monitor, Evaluate and Assess and "04" is the number of the objective).

The description of all objectives follows the idea, that for the establishing and tailoring the objective the enterprise needs to take in account a number of components (e.g., processes, organizational structures, information and others). Each objective always relates to one process. Focusing on the chosen objective MEA04 Managed Assurance the related process - Assurance Management - is described by the set of practices and activities to achieve objective and produce a set of outputs that support achievement of overall IT-related goals. COBIT Core Model assigns capability levels to all process activities, enabling clear definition of the process and required activities for achieving the different capability levels. The process within objective can operate at various capability levels, ranging from 0 to 5. The capability level is a measure of how well a process is implementing and performing. Capability levels follow the CMMI 2.0 process capability scheme.

The information received from the COBIT Core Model description of the process Assurance Management is summarized in Table 3.

Table 3: List of activities of Assurance Management process assigned to capability level 2

Practice	Activities with the assigned capability level 2	F	L	P	N
1. Ensure assurance independence	Establish adherence to applicable codes of ethics and standards				
	Establish independence of assurance providers				
	Establish competency and qualification of assurance providers.				
2.Develop risk-based planning	Understand the enterprise strategy and priorities.				
	Understand the internal context of the enterprise				
	Understand the external context of the enterprise				
3. Determine the objectives	Define the assurance objective of the assurance initiative by identifying the stakeholders of the assurance initiative and their interests.				
	Agree on the high-level objectives and the organizational boundaries of the assurance engagement				
4. Define the scope	Define all governance components in scope: principles, policies and frameworks; processes; organizational structures; culture, ethics and behavior; information; services, infrastructure and applications; people, skills and competences				
5. Define the work program	Define detailed steps for collecting and evaluating information from management controls within scope				
	Understand the context of the management objectives and the supporting management controls that are put in place				
	Understand all stakeholders and their interests				
6.Execute the assurance initiative focusing on design	Refine the understanding of the IT assurance subject.				
	Refine the scope of the IT assurance subject.				
7.Execute the assurance initiative focusing on operating effectiveness	N/A				
8.Report and follow up	Document the impact of control weaknesses				
	Communicate with management during execution of the initiative so there is a clear understanding of the work performed and agreement on and acceptance of the preliminary findings and recommendations				

9.Follow up on recommendations	Agree on and implement internally, within the organization, the necessary actions that need to be taken to resolve identified weaknesses and gaps				
	Follow up, within the organization, to determine whether corrective actions were taken and internal control weaknesses were resolved				

Table 3 summarizes the activities demanded for the achievement of capability level 2. Similar tables can be received for capability levels 3 – 4. Capability levels 1 and 5 are in case of the chosen process omitted. The likely reason is that achieving capability level 1 for IA function is meaningless and achieving level 5 unrealistic. In the first column are listed the process practices. The process assumes the implementation of 9 practices that cover the entire life cycle of the assurance function. The second column summarizes the activities of the practices that are assigned to the capability level 2. Notice, that the Practice 7 does not have any assigned activity for this level. The last four columns are the support for assessor as they represent less formal method leading to independent ratings based on CMMI recommended scales (ISACA, 2018, p 39):

- Fully - The capability level is achieved for more than 85 percent. (This remains a judgment call, but it can be substantiated by the examination or assessment of the components of the enabler, such as process activities, process goals or organizational structure good practices.)

- Largely - The capability level is achieved between 50 percent and 85 percent.

- Partially - The capability level is achieved between 15 percent and 50 percent.

- Not - The capability level is achieved less than 15 percent.

The output of the IA assessment is the completed table. The conclusion from the received data is the next (Souza J.,2019) :

- If all level 2 activities in every practice have been rated L or F, this process, at least, has to attend the requirements of level 2.

- If any level 2 activities in all practices of the process have been rated N or P, then:
  - Evaluate whether to reach the goals defined for this process, it is necessary, somehow, to achieve its purpose.
  - If it is necessary, then a capability level of 1 should be the target for the process.
  - Otherwise, the process should be set aside.

5. Conclusion

Currently there exist the whole set of methods and models supporting the assessment of the IA performance. IA assessment is the part of quality assurance and improvement program (QAIP) regulated by IPPF Standards. Table summarizes the main differences between the identified models.

	Traditional model	Maturity model	Capability model
Main aim	To be compliant to standards	To have all IA activities at the same level of maturity	To provide continuous improvement of IA activities
Main components	Separate standards	IA activities based on standards	Separate processes/practices/activities according assurance lifecycle
Flexibility	No context	Limited space for adding context	Space for adding context
Recommended type of assessment	Internal self assessment	Internal self assessment (in combination with external assessment)	External assessment of the IA activities at the enterprise wide level
Periodicity	Ongoing monitoring	Each year	Once a five years

Users	Auditors at the engagement level	IA managers, departmental level	board, audit committee and management of the organization

Table 4: Overview of the differences between the various models of IA assessment

Capability model based on COBIT 2019 is a new option and future developments will show its usability.

6. References

EU (2013), Public Internal Control Systems in the European Union, Quality Assurance for Internal Audit, Discussion paper No. 3

IIAa (2019), Mandatory Guidance, Definition of internal auditing, theiia.org

IIAb (2017), The Institute of Internal Auditors Quality Assessment Manual for the Internal Audit Activity, 2017, theiia.org

IIAc (2020), Position Paper, Internal Auditing´s Role in Corporate Governance, theiia.og

IIAd (2017), Internal Audit Maturity Model, available at https://na.theiia.org/services/quality/Public_Documents/maturity%20model.pdf

IPPF (2013), Practice Guide: Selecting, Using, and Creating Maturity Models: A Tool for Assurance and Consulting Engagements, July 2013, available https://na.theiia.org/standards-guidance/Public%20Documents/PG%20Maturity%20Models.pdf

ISACA (2018), COBIT 2019 Framework: Introduction and methodology, ISBN-978-1-60420-763-7, pp 37

Pempal (2016) Pempal IACOP, Quality Assessment Guide for Public Sector Internal Audit

Schneider (2010), ASSESSMENT OF INTERNAL AUDITING BY AUDIT COMMITTEES, Georgia Institute of TechnologyAcademy of Accounting and Financial Studies Journal, Volume 14, Number 2, 2010, p 20

Souza J. (2019) Defining Target Capability Levels in COBIT 2019: A Proposal for Refinement, available at https://www.researchgate.net/publication/332531351_Defining_Target_Capability_Levels_in_COBIT_2019_A_Proposal_for_Refinement

THE EFFECTS OF NEGATIVE EQUITY ON THE BUSINESS PERFORMANCE IN SLOVAKIA

Katarína Tasáryová, Renáta Pakšiová

Faculty of Economic Informatics
University of Economics in Bratislava
k.tasaryova@gmail.com, renata.paksiova@euba.sk

Keywords

performance, equity, negative equity, financial analysis, Altman's Z-score

Abstract

In order for individual companies to be successful, meet their objectives and continually increase competitiveness, it is necessary to measure and evaluate their performance using appropriately selected performance indicators. The performance of a business is closely linked, in particular, to the capital that the business can use without repayment in the long run, i.e. own resources. Own resources in accounting are equity, which is the difference between the assets and liabilities of the business. If the value of the liabilities exceeds the value of the assets or if the value of the assets is lower than the liabilities of the business, the negative equity arises. Negative equity is currently a common phenomenon encountered by many companies. The aim of the paper is to analyze the impact of negative equity on the business performance in Slovakia. Using correlation analysis, we identify the correlation between negative equity and the business performance, in particular the direct dependency of the bankruptcy zone according to Altman's Z-score against negative equity, in order to indicate possible ways to improve the financial position of the own resources of the business and prevent its liquidation.

1. Introduction

One of the main intentions of managers in terms of a financial management of the business is to constantly increase the business performance. The business performance is closely related to the existence of the business itself. It is about valuing the capital invested in the business and subsequently increasing the value of the business. The purpose of performance management is to achieve the set goal through continuous improvement of all key characteristics of the business, and thus increase its performance at all levels of management (Melecký and Staníčková, 2018) and vis-à-vis the highly competitive market (MacGregor Pelikánová, 2019). According to Mardones and Cuneo (2019), business performance is a key issue for investors, shareholders and the functioning of the economy (p. 1). The business performance emphasizes variables directly related to financial statements. In view of the above, the performance of companies can be analyzed according to various criteria, such as profitability, dividend growth, sales turnover, capital employed and so on. The business performance is the ability of a business to achieve and manage business resources in a variety of ways to develop a competitive advantage (Yudhanti and Shanti, 2011, p. 60). Wagner (2009) understands performance as a characteristic describing the way or course in which a given subject carries out an activity on the basis of similarity to the reference method of performing that activity, and the interpretation of this characteristic presupposes the ability to compare the examined

and reference phenomenon (p. 17). According to Almajali et al. (2012) there are different measures of a financial performance of a business, for example, the return on sales informs how much the business earns in connection with its sales, return on assets explains the ability of a business to use its assets and return on equity informs about what investors will get for their investments (p. 270). Liargovas & Skandalis (2010) argue that there is still a debate among various disciplines on the correct measurement of business performance as well as the identification of factors that affect it (p. 184). There are many factors influencing the business performance in terms of capital structure, the most important being the structure of owners (Wei et al., 2020, p.71).

The literature usually distinguishes between the two types of a business performance, namely financial or economic and innovative performance. According to Donkor et al. (2018), the financial performance is measured by means of seven items, including sales volume, profit levels, return on sales, revenue growth, return on investment, market share and profitability growth (p. 244). The innovative performance is expressed by means of expenses, patents or the percentage of innovative sales (Hagedorn and Cloodt, 2003, p. 1367). Although both types of the performance are closely related (Damanpou and Evan, 1984, p. 392), the literature uses them as separate concepts or focuses on only one of them (Knoben and Oerlemans, 2006, p. 4). The difference in the development of performance indicators is smaller for innovative businesses (Bockova and Zizlavsky, 2016, p.173). Integrated performance measurement systems belong to an important strategic management accounting techniques by businesses (Petera and Soliakova, 2020).

The business performance is closely linked, in particular, to the capital that the business can use without the obligation to repay in the long term, in this case its own resources. When a business is set up, it is capital raised from investors. In connection with equity invested in the business, it is important to identify whether the investors are domestic or foreign, especially because of the risk of outflow capital generated by the business in the form of profit abroad. Foreign capital can be understood in two senses: either it is a capital invested by foreign investors in Slovakia or it is a capital invested by Slovak investors abroad. In the first sense, in connection with the evaluation of the financial situation of the business, it may be related to the incurrence of liabilities in the payment of profit shares. The business can create its own resources during its existence, for example from the achieved profit. Own resources represent equity in the accounting, which is indirectly defined in the Slovak legislation within the section 2(2)c) of Act No. 431/2002 Coll. on accounting, as amended, as the difference between assets and liabilities of the business. The main component of equity is the capital, which in Slovakia must be created by limited liability companies, joint stock companies and simple companies for shares (Act No. 513/1991 Coll., the Commercial Code, as amended, section 58(2)). In addition to the capital, the equity of business entities consists of the following items: capital funds, valuation differences, funds created from profit, profit or loss for the accounting period after tax and profit or loss from previous years. According to Valaskova et al. (2019), the volume of equity decisively determines the long-term survival of the business (p. 1). If the equity reaches a negative value, it means that the amount of liabilities exceeds the amount of assets, or that the business's assets are financed mainly through external sources representing liabilities (Srnišová, 2018, p. 382).

The evaluation of the business financial performance is carried out by means of a financial analysis. Cisko & Kliešitk (2013) define a financial analysis as a tool for diagnosing economic systems, enabling, on the one hand, to reveal the influence of economic and non-economic factors and, on the other hand, to estimate their future development (p. 573). The main tasks of financial analysis include reporting the property and financial situation of the business and preparing documents for managerial decision-making in the internal environment. A business that can achieve the required rates of capital appreciation by investors in the long run can be considered a financially health business, a prosperous business. On the contrary, a business whose own resources are reduced to negative values gets into problems in terms of its long-term existence, which can lead to bankruptcy and business termination.

For this reason, it is important to address the issue of analyzing the impact of negative equity on the business performance in Slovakia in time series.

According to some studies, there is a positive correlation between capital concentration and business performance. Pedersen & Thomsen (1999) examined the causes and effects of the concentration of ownership among the largest businesses in 12 European countries (p. 367). They found that the concentration of ownership did not have a direct effect on the return on invested capital and that businesses benefit from the diversification of ownership. This could have important consequences for large European businesses, especially those seeking to grow rapidly At the same time, large owners may have a better situation in case of problems with information and uncertainty of investors in the area of assessing the situation in the business. Ke & Isaac, (2007) examined the relationship between the ownership structure and the performance of listed real estate companies in China in 2000-2002 (p. 9). The research has shown that in the real estate sector, government shares have a significant impact on the business performance in China. Vătavu (2015) examined the relationship between the capital structure and the financial performance of 196 Romanian businesses listed on the Bucharest Stock Exchange and operating in the manufacturing sector in 2003-2010 (p. 1314). The research has shown that in 2003-2010, the most profitable were those manufacturing businesses that maintained a high share of equity in their capital mix and avoided borrowed resources. Equity has a positive effect on performance indicators, while a total debt and a short-term debt have negative relationships with ROA and ROE. Makhija & Spiro (2000) conducted research on a sample of 988 newly privatized Czech businesses, based on which they found a positive correlation between share ownership and business value (p. 1). Our current research into the impact of negative equity on business performance is important, especially because of the expected impending economic crisis, when such cases will multiply.

2. Data and methodology

Quantitative research carried out on a sample of economic information of businesses established in Slovakia for the period of 2014-2018 aims to identify the impact of a negative equity on the performance of these businesses. Information relevant to the research is provided in the financial statements of the examined businesses, in particular in the balance sheet and profit and loss account for the relevant accounting periods. Businesses record their financial statements in a publicly available Internet source of information - the register of financial statements (www.ruz.sk). The database of necessary financial information of the examined businesses for the period of 2014-2018 is processed on the basis of completed data from www.finstat.sk, which regularly synchronizes information about businesses with the data published in the register of financial statements and had needed information for our research.

Subjects of the research are businesses established in Slovakia that have their published financial statements for the accounting period of 2014-2018. The research sample does not include those businesses that are in bankruptcy or restructuring proceedings, and those in which we have identified inconsistencies in reporting. Specifically, in 2014 there were 181,153 businesses examined, in 2015 the number of businesses increased to 184,114, in 2016 the subject of research was 191,240 businesses, in 2017 the number increased to 201,075 businesses and in 2018 we saw the highest number of businesses established in the Slovak Republic, specifically 211,824. Based on information from the financial statements of each business, we analyzed the structure of equity of individual businesses for the period of 2014-2018. Subsequently, we assessed the performance of businesses that showed the negative equity during the period under review using the performance indicators, which means we excluded businesses with zero and positive equity from the research. Using correlation analysis we assessed the relationship between the negative equity and the performance of

the examined businesses and the results found in terms of the possibility of their use in improving the financial situation, especially the path from the bankruptcy zone.

2.1. Model and Data

When evaluating the business performance we use the Altman Z-score, also known as the Altman index of financial health, which was created on the basis of a discriminatory analysis, by means of which Professor Altman tried to find indicators best distinguishing prosperous businesses from the non-prosperous ones leading to bankruptcy (Kubíčková and Jindřichovská, 2015, p. 207). Altman's Z-score (1) is expressed by the following quantities (Holečková, 2008, p. 195): net working capital (NWC), total assets (TA), return on net assets (ROA), earnings before interest and taxes (EBIT), book value of equity (BVE), liabilities (L), sales (S).

$$Z = 0.717 * \frac{NWC}{TA} + 0.847 * ROA + 3.107 * \frac{EBIT}{TA} + 0.42 * \frac{BVE}{L} + 0.998 * \frac{S}{TA} \qquad (1)$$

The resulting value of Altman's Z-score according to Růčková (2011, p. 72) indicates a satisfactory financial situation, significant financial problems or the value is in the gray zone (Table 1).

Table 1 – Characteristics of Altman Z-score values

Value of Altman's Z-score	Characteristic
Z > 2.99	prosperity zone (satisfactory financial situation)
1.2 < Z ≤ 2.9	gray zone (it is not possible to clearly assign if the business is prosperous or not)
Z < 1.2	bankruptcy zone (relatively significant financial problems)

3. Results

In the monitored period of 2014-2018, the structure of the equity of businesses established in the Slovak Republic varied, with an average of 131,805 companies showing positive equity during the whole period. On average, 56,742 businesses reported negative equity over the entire monitored period. In 2016, the smallest number of businesses showed negative equity (Figure 1).
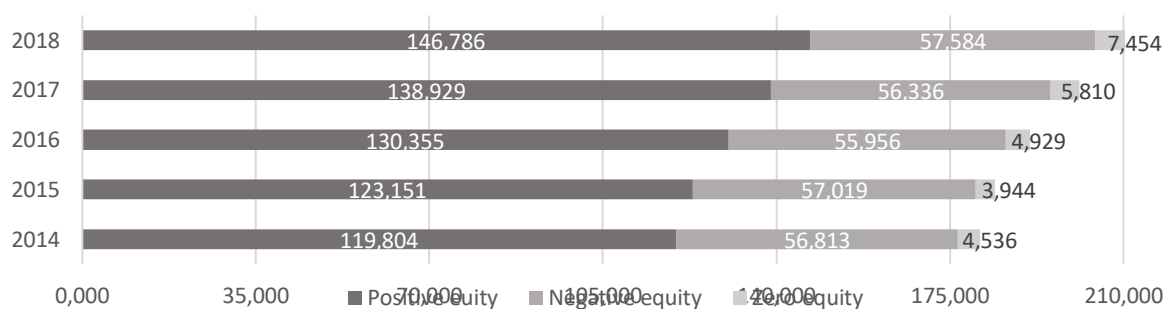
Figure 1 – Overview of equity value of businesses established in SR in the review period

Based on the Altman Z-score, out of the businesses with the negative equity in the monitored period, on average 69% have significant financial problems, putting these businesses in a zone of bankruptcy. This percentage corresponds to approximately 39,182 businesses. Only an average of 18% of businesses are in the prosperity zone, which means that these businesses have a satisfactory financial situation. This is an average of 10,066 businesses. Given that these businesses show the negative equity as well, the individual indicators needed to calculate the Altman Z-score are subject to further analysis (Figure 3). For the remaining 13% of businesses we cannot determine with a certainty

whether they have a satisfactory financial situation or not, because these businesses are in the gray zone (Figure 2).
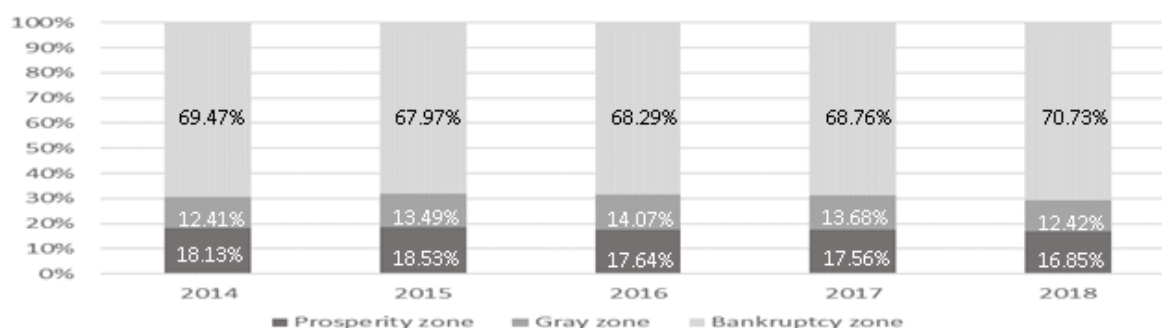
Figure 2 – Development of business financial performance with the negative equity
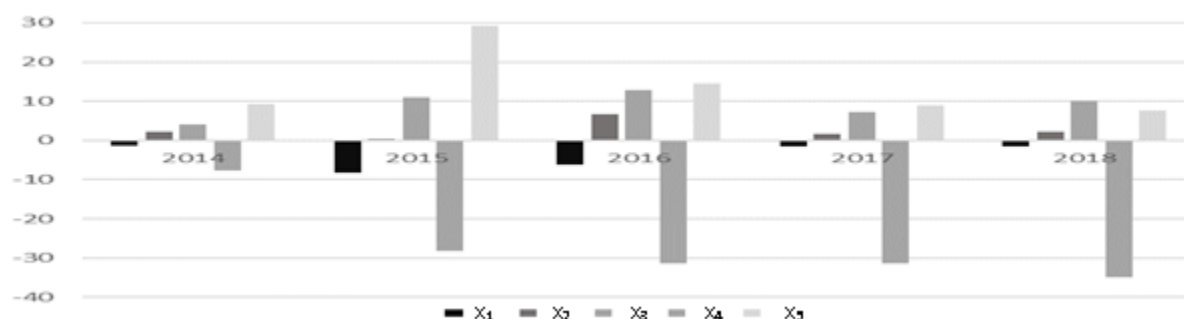
Figure 3 – Average values of indicators needed to calculate the Altman Z-score

When calculating Altman's Z-score, the relationship between the equity and liabilities (x_4) is also considered. Given that the businesses that report the negative equity are the subject of analysis, the value of x_4 is negative in the reporting periods. However, this did not have an absolute impact on the performance of the examined businesses, as the chosen performance evaluation indicator (in our case the Altman Z-score) also takes into account other items than equity itself. The variable x_1 represents the ratio of net working capital to total assets. The value of that variable was negative at all times but did not strong affect the performance of the undertakings. Variable x_2 represents the return on assets, which is the ratio of accounting profit to the total assets of the business. As the negative equity in the period under review may still be reported by businesses that generated the negative equity in previous periods and have already made a profit in the period under review and their negative equity is declining in absolute terms, the average value of the examined businesses is positive, meaning that the financial situation of the undertakings concerned is already improving. The same corresponds to the average amount of variable x_3, which represents the ratio of profit before tax and interest to total assets, and the last variable x_5, which gives the ratio of sales to total assets. The variables x_2, x_3 and x_5 were positive in the examined businesses in the monitored periods and significantly influenced the calculated value of Altman Z-score, therefore these businesses can be considered as prosperous according the criteria. In a correlation analysis carried out on a sample of businesses reporting the negative equity for the relevant periods, a correlation coefficient of 0.9 was identified between the negative equity and the bankruptcy zone. It can be stated from the above that there is a strong dependence between these elements. If the businesses are in the bankruptcy zone, the negative equity will negatively affect their existence, as they will not have any own funds to meet their obligations. A completely opposite situation occurred in case of the negative equity and the gray zone, where a correlation coefficient of -0.7 was identified. It follows from the above result that there is an indirect dependence between the examined elements. The last elements examined were the negative equity and prosperity zone, among which a correlation coefficient of 0.08 was identified. This result means

that there is a minimal dependency between the elements, so it can be concluded that prosperous businesses will not be significantly affected by the negative equity. These results were also confirmed by the graphical dependence (Figure 4), from which we can see that the negative equity and the bankruptcy zone are almost linear, while the prosperity zone and the gray zone have a different course.
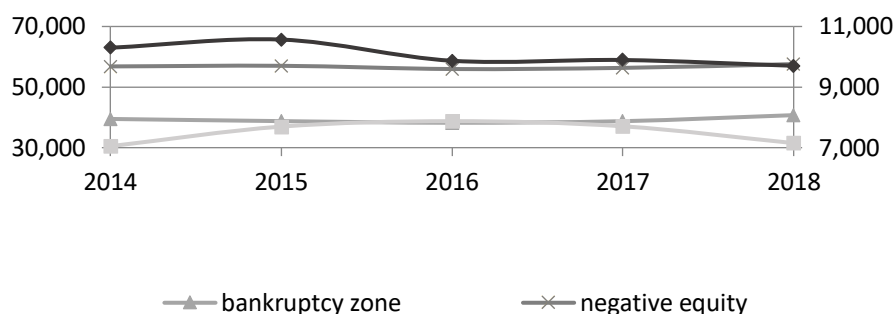
Figure 4 – Analysis of the relationship between the negative equity and business performance

4. Conclusion

The equity is an own resource of the business that has been available for a long time. When a business is established, the equity consists of deposits made by individual shareholders in the business. The most important component of the equity is the capital representing the monetary expression of the sum of monetary and non-monetary contributions of all shareholders to the business (Commercial Code, section 58 (1)). A joint-stock business, a limited liability company and a simple company for shares are required by law to create capital, the amount of which is entered in the Commercial Register.

If the value of the liabilities exceeds the value of the assets, a negative equity is created. At present, the negative equity is a frequent phenomenon encountered by many businesses. The aim of the paper was to analyze the impact of the negative equity on the business performance in Slovakia. The research focused on businesses established in Slovakia in the period of 2014-2018. The results of the analysis show that 69% of businesses are in the bankruptcy zone, which means that they have significant financial problems. Only 18% of businesses with the negative equity are in the prosperity zone, which means that these businesses, despite their negative equity, have an improving financial situation. The rest of the businesses are in the gray zone, where it is impossible to determine if the business is prosperous or not. The correlation analysis carried out showed that there is a strong dependence between the negative equity and the bankruptcy zone, so that the negative equity will significantly affect not only the business performance in the bankruptcy zone but also their existence. In case of businesses located in the gray zone, there is an indirect dependence, which results from the fact, that it is not possible to clearly determine if businesses located in the gray zone are prosperous or not. For businesses located in the prosperity zone the negative equity does not have a material effect on the performance of these businesses, because a minimal dependence has been identified between the examined elements.

Despite the results obtained, the direct correlation between the negative equity and the bankruptcy band according to the Altman Z-score, the examined businesses with the negative equity, which are in the bankruptcy zone, may not automatically cease to operate. The best solution for businesses in the bankruptcy zone is to increase the equity by making a profit in subsequent accounting periods, when businesses can quickly reach a zone of prosperity, as we have shown in research. In order to achieve profits in a difficult financial situation of the negative equity, it is important for the business to ensure a high degree of efficiency in the use of resources, cost-effectiveness in achieving revenues

(Turečková and Nevima, 2020). The use of innovative performance is also one of the important ways in the period of scientific and technical progress to improve the financial situation and to bridge the period of crisis, which can also be caused by external conditions, such as a result of a global pandemic, a mortgage crisis or a crisis of natural resources.

Except the most common additional owners' deposits to businesses Srnišová (2018) mentions three ways of increasing equity, namely by forgiving the liability, by contributing to the capital fund or by capitalizing the liability in the capital (p. 384). Increasing equity by forgiving the liability is administratively undemanding for businesses. A written document on the waiver of debt serves as a basis for an accounting. The disadvantage of this method is the burden of income tax on businesses, so it is recommended to use this method in cases where the business has the possibility of amortization of tax losses from the previous periods. When the equity is increased by a contribution to the capital fund, the liability to the shareholder is reclassified. The contribution to the capital fund does not affect the profit or loss. There is no change in the total amount of assets or liabilities, but there is a change in the structure of liabilities. Advantages of this method of increasing the equity include also administrative undemanding, but also the possibility of paying back the contribution to the capital fund back to the shareholder. The last option to increase the equity is to capitalize liabilities, which can be an interesting choice for a creditor who is not a partner of a debtor and has the opportunity to become one of the owners by capitalizing his claim, thus gaining all the owner-related benefits related to the asset claim and decision-making rights of the business. This method is the least flexible compared to the previous ones, because the business is bound by law in this case.

5. Acknowledgements

6. References

Act No. 431/2002 Coll. on accounting, as amended.

Act No. 513/1991 Coll., Commercial Code, as amended.

Almajali, Y. A., Alamro, S. H., & Al-Soub, Y. Z. (2012). Factors Affecting the Financial Performance of Jordanian Insurance Companies Listed at Amman Stock Exchange. Journal of Management Research, 4 (2), 266-289.

Bockova. N. & Zizlavsky, O. (2016). The Innovation and Financial Performance of a Company: A Study from Czech Manufacturing Industry. Transformations in Business and Economics 15(3), 156-175.

Cisko, Š. & Klieštik, T. (2013). Finančný manžment podniku II, Žilina, EDIS, ISBN 978-80-554-0684-8.

Damanpour, F., & Evan, W. M. (1984). Organizational Innovation and Performance: The Problem of "Organizational Lag." Administrative Science Quarterly, 29(3), 392.

Donkor, J., Donkor, G. N. A., Kankam-Kwarteng, C. & Aidoo, E. (2018). Innovative capability, strategic goals and financial performance of SMEs in Ghana. Asia Pacific Journal of Innovation and Entrepreneurship, 12 (2), 238-254.

Finstat. (2020). Dataset of financial information of companies from SR for 2014 – 2018. (Available at www. finstat.sk).

Hagedoorn, J. and. Cloodt, M, (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? Research Policy, 32(8), 1365-1379.

Holečková, J. (2008). Finanční analýza firmy. Praha, ASPI, ISBN 978-80-7357-392-8.

Ke, Q. & Isaac, D. (2007). Ownership structure and corporate performance: empirical evidence of China's listed property companies. Journal of Financial Management of Property and Construction, 12(1), 3-10.

Knoben, J. & Oerlemans, L. (2005). The effects of firm relocation on firm performance – A literature review, , European Regional Science Association, 1-25.

Kubíčková, D. & Jindřichovská, I. (2015). Finanční analýza a hodnocení výkonnosti firem, Bratislava, C. H. Beck, ISBN 98-80-740-0538-1.

Liargovas, P., & Skandalis, K. (2010). Factor affecting firms' financial performance: The Case of Greece. Athens: University of Peloponnese Press. Global Business and Management Research: An International Journal, 2(2-3), 184-197.

Makhija Anil K. & Spiro, M. (2005). Ownership Structure as a Determinant of Firm Value: Evidence from Newly Privatized Czech Firms. Financial Review 35(3), 1-31.

Mardones, J. G. & Cuneo, G. R. (2019). Capital structure and performance in Latin American companies. Economic Research-Ekonomska Istraživanja, 1-19.

MacGregor Pelikánová, R. (2019). Harmonization of the protection against misleading commercial practices: ongoing divergences in Central European countries. Oeconomia Copernicana, 10(2), 239–252.

Melecký, L. & Staníčková, M. (2018). Complexity of the Project Cycle Management and Logical Framework Approach: Challenges or Standards in the EU Case? In Staníčková, M., L. Melecký, E. Kovářová and K. Dvoroková (eds.). Proceedings of the 4th International Conference on European Integration 2018. Ostrava: VŠB - Technical University of Ostrava, 2018, pp. 1035-1042. ISBN 978-80-248-4169-4. ISSN 2571-029X.

Pedersen, T. & Thomsen, S. (1999). Economic and Systemic Explanations of Ownership Concentration among Europe's Largest Companies. International Journal of the Economics of Business, 6(3), 367-381.

Petera, P & Soljakova, L. (2020). Use of strategic management accounting techniques by companies in the Czech Republic. ECONOMIC RESEARCH-EKONOMSKA ISTRAZIVANJA, 33(1), 46-67.

Register FS. (2019). Register of financial statements 2014 – 2018. (Available at http://www.registeruz.sk/cruz-public).

Růčková, P. (2011). Finanční analýza: metody, ukazatele, využití v praxi, Praha, Grada, ISBN 078-80-247-3916-8.

Srnišová, Petra. (2018). Záporné vlastné imanie ako indikátor existencie transakcií medzi spriaznenými osobami: Negative Equity as an Indicator of the Existence of Transactions Between Related Parties. MMK 2018: recenzovaný sborník příspěvků z mezinárodní Masarykovy konference pro doktorandy a mladé vědecké pracovníky, ročník IX., 17. – 21. prosince 2018, Hradec Králové, Česká republika. Hradec Králové: MAGNANIMITAS, 2018, 381-386.

Turečková, K. & Nevima, J. (2020). The Cost Benefit Analysis for the Concept of a Smart City: How to Measure the Efficiency of Smart Solutions? Sustainability, 12(7), 1-17.

Valaskova, K., Lazaroiu, G., Olah, J., Siekelova, A. & Lancova, B. (2019). How capital structure affects business valuation: A case study of Slovakia. Central European Business Review, 8(3), 1-17.

Vătavu, S. (2015). The impact of capital structure on financial performance in Romanian listed companies. Procedia Economics and Finance 32, 1314-1322.

Wagner, J. 2009. Měření výkonnosti. Jak měřit, vyhodnocovat a využívat informace o podnikové výkonnosti. Praha, GRADA Publishing, ISBN 978-80-247-6739-0.

Wei X., Che H., Sheng N., Hsiao CY-L, Tong Q. & Yan G. (2020). Research on the Development Status of China's Renewable Energy Industry - The Impact of Capital Structure on Company Performance. Frontiers in Energy Research, 8 (71), 1-10.

Yudhanti, C. B. H. & Shanti, C. J. (2011). Intellectual Capital dan Ukuran Fundamental Kinerja Keuangan Perusahaan, Jurnal Akuntansi dan Keuangan, 13(2), 57-66.

CONTROLLING IN RELATION TO DIVISION OF ENTERPRISES AND COMPANY SIZE

Miroslava Vlčková

Faculty of Economics
University of South Bohemia
Czech Republic
mvlckova02@ef.jcu.cz

Keywords

Controlling, Divisions of enterprises, Enterprise size, Mann-Whitney U test

Abstract

The implementation of controlling elements is nowadays essential for most companies. There are even companies that take the controlling and controlling department as an obviousness without which they would not be able to manage the company correctly and responsibly. This paper focuses on the analysis of enterprises in terms of their size given by the number of employees and the selection of enterprises by industry in relation to controlling. In total, almost 600 companies with 10 to 250 employees were analyzed. It was found that the implementation of controlling in the companies was directly related to the number of employees. Regarding individual sectors, companies were first divided into nine categories in relation to CZ-NACE and then the analysis of controlling in divisions of enterprises was performed. It has been found that the size of enterprises in the administrative and support activities; professional, scientific and technical activities; wholesale and retail and manufacturing is directly linked to the implementation of controlling (at a significance level 0.05) and the size of enterprises in the transport and storage sectors; information and communication activities; services; construction industry and agriculture, forestry, fishing has no direct link to controlling.

1. Introduction

Controlling ensures that the organization's resources are used efficiently and effectively to achieve predetermined goals. Many factors will affect the implementation of controlling elements such as size of the enterprise, number of employees, sector in which the enterprises do the business, financial indicators such as for e. g. amount of assets and liabilities, value of liquidity, profitability, ways of financing. This paper is focused on the size of the company analyzed according to the number of employees, further on the influence of divisions of enterprises in relation to controlling. Obviously, other approach to controlling will have small enterprises and different approach to controlling will be implemented by medium and macro enterprises. This is also confirmed by López & Hiebl (2015) who say that the use of controlling is different in SMEs compared to larger entities.

Controlling verifies that everything is in the line with the organization's plans, guidelines and established policies. This coordination ensures effective using of organizational resources to achieve organizational plans. It measures the variations of actual power from the standard. It also identifies the causes of these variations and helps in taking corrective action (Cornel, & Lavinia-Maria, 2012).

Also Fibírová (2003) deals with the size of the company in connection with controlling. She says that controlling elements such as comparison and subsequent evaluation of internal data are usually evaluated outside the accounts in smaller businesses, while larger companies require a larger system of management accounting and controlling information.

2. Literature review

Controlling is understood as a method for increasing the efficiency of the management system by constantly and systematically comparing the facts with a predetermined goal, evaluating, identifying the causes of differences, proposing corrective measures, etc. Controlling can be defined as:

"... a management tool to coordinate the planning, control, and provision of an information base to improve business results." (Horváth, 2003)

It follows from this definition that the aim of providing controlling information is to comprehensively capture events so that an appropriate information base for management is established. It can also be concluded that the information expressed in financial accounting does not reflect other information relevant to management; it is merely a value expression of certain facts.

Primary purpose of controlling is coordinate management systems for cover information and also to ensure harmonization within and outside the company. Managers need an information system that will identify and solve problems, such as the possibility of cost overruns or the inability to implement the plans for future (Drury, 2015).

Controlling models can be understood here as structured management tools that are considered particularly important for success (Magretta, 2002). This statement is followed by Wirtz et al. (2016), who define four basic areas of research in this field, namely innovation, change and development, performance and control and design.

Controlling proposes systematic monitoring of enterprise resource potential not only in real time, but is also included in the overall management system through strategic (predictive) management. It is used to identify potential business capabilities and to manage the potential of business resources, namely control (Čebukina, et al. 2013).

Importance of controlling implementation is that this system helps to company in (Stoner, Freeman, & Gilbert, 1998):

- the controlling function measures progress and brings to light the variations, if there are any, and indicates corrective action to be achieved the goals,

- the controlling system enables management to verify whether the standards set are accurate and objective and helps to review and revise the standards in light of such changes,

- the controlling ensures that resources are used in the most effective and efficient manner,

- the controlling system ensures that employees know well in advance what they are expected to do and it motivates them and helps them to give better performance,

- the controlling provides direction to all activities and efforts for achieving organizational goals. each department and employee is governed by predetermined standards which are well coordinated with one another. this ensures that overall organizational objectives are accomplished.

It can be seen from the above that controlling has an important role in the management of the company. As Reichmann (1997) says that the basic objectives of controlling are goals, that determine the reasons for planning and design management systems. These goals are derived from

organization's main goals. The objectives are primarily in the form of profit, productivity and liquidity.

Economy helps managers to get the best resources of appropriate quantity and quality at minimum level of costs. Measurement of the economy is based on a comparison the actual costs with the planned costs. There are two ways how to increase economy – by the absolute savings and by the relative savings (Král, 2010).

Controlling can be viewed from several perspectives or areas. Simkin (2010) examines the barriers in implementing marketing plans within controlling as analyzed over several years in the field of marketing planning. It also deals with controlling activities that try to overcome the barriers and lead to other benefits. Controlling is very close to managerial accounting. Managerial accounting is one of the important parts of financial controlling. Research by Bui, Tu Le, & Nguyen (2020) found that the level and content of managerial accounting in individual companies varied according to size, length of business and business sector. Small companies often use common financial accounting methods, while large companies used management accounting methods such as activity-based costing, activity-based budgeting, decision analysis, target costing and strategic planning.

Vlčková (2018) concerned with the management accounting survey and seeks to identify the financial variables, whose amount is different for businesses that do not have implemented management accounting from companies with implemented management accounting. The adoption of elements of financial controlling are fundamental events of new and emerging companies. In connection with this, however, there is an increase in the number of employees of the company when introducing controlling elements. On the other hand, their faster roll-out is also associated with fast-growing companies (Davila & Foster, 2005).

3. Methodology

The aim of this paper is to analyze data obtained from 592 companies from the Czech Republic and their relationship to controlling. These are the enterprises with a number of employees ranging from 10 to 250. The data were collected by means of questionnaire surveys, when a proportional sample of 44,000 enterprises in terms of the business sector was created, subsequently, in 2018, data from 592 companies were obtained. The enterprises were divided into 9 categories according to the business sector using the CZ-NACE methodology. These categories are: administrative and support activities, transport and storage, information and communication activities, professional, scientific and technical activities, services, construction industry, wholesale and retail, agriculture, forestry, fishing, manufacturing.

At the same time, the companies were divided according to whether they have implemented controlling elements in the company, where it was proved that 323 companies had implemented controlling, compared to 269 companies without implemented controlling. Subsequently, the individual relationships of these factors were analyzed and the relationships between the company size, divisions of enterprises and controlling were analyzed.

As a statistical test, Mann-Whitney U test was used. This test is used to evaluate unpaired experiments when comparing two different samples. It was tested the hypothesis that two variables have the same probability distribution. At the same time, these variables may not correspond to Gaussian normal distribution, it is sufficient to assume that they are continuous. The test involves the calculation of a statistic whose distribution under the null hypothesis is known. The null hypothesis asserts that the medians of the two samples are identical (Freund, Wilson, Mohr, 2010; Budíková, Králová, Maroš, 2010)). Statistically, it has always been tested at a significance level of 0.05 where zero hypothesis H0: $\mu_1 - \mu_2 = 0$ against alternative hypothesis HA: $\mu_1 - \mu_2 \neq 0$. U is then given by:

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2}, \tag{1}$$

where n_1 is the sample size for sample 1, and R_1 is the sum of the ranks in sample 1. An equally valid formula for U is:

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \tag{2}$$

The smaller value of U1 and U2 is the one used when consulting significance tables. The sum of the two values is given by:

$$U_1 + U_2 = R_1 - \frac{n_1(n_1+1)}{2} + R_2 - \frac{n_2(n_2+1)}{2} \tag{3}$$

Knowing that $R_1 + R_2 = \frac{N(N+1)}{2}$ and $N = n_1 + n_2$, and doing some algebra, we find that the sum is $U_1 + U_2 = n_1 n_2$.

4. Results and discussion

The main goal of this paper is to evaluate the relationship between the size of the company (evaluated according to the number of employees) in comparison with the implementation of controlling. A partial goal is to evaluate the relationship between the divisions of enterprises by industry in comparison with implementation of controlling. Table 1 shows the breakdown of enterprises according to whether or not they have implemented controlling and at the same time a breakdown of these enterprises according to business areas. The survey shows that more than half of the companies have implemented controlling. Controlling is most often implemented in the divisions of the manufacturing industry. On the other hand, there is a surprising result in the division of enterprise agriculture, forestry and fishing where only one quarter of the monitored enterprises have implemented controlling.

Table 1: Divisions of enterprises with implemented controlling

Division of enterprise	Controlling YES	Controlling NO	Total	Share of YES
Administrative and support activities	12	14	26	46%
Transport and storage	13	12	25	52%
Information and communication activities	30	21	51	59%
Professional, scientific and tech. activities	33	38	71	46%
Services	13	16	29	45%
Construction industry	39	26	65	60%
Wholesale and retail	60	46	106	57%
Agriculture, forestry, fishing	7	21	28	25%
Manufacturing	116	75	191	61%
All divisions of enterprises	323	269	592	55%

Source: Own research

The next step was a statistical analysis of the relationship between the size of the company and divisions of enterprises compared to the implementation of controlling. It was used a two-sample Mann-Whitney U test at the selected significance level α = 0,05. The hypotheses H0 = x0,50 – y0,50 = 0 were tested, where it is assumed that the influence of implementation of controlling in these enterprises is the same in both groups and the hypothesis HA= x0,50 > y0,50, which assumes that the enterprises without implementation of controlling will have less employees.

In both hypotheses, X represents the companies with implemented controlling while Y represents the enterprises without implementation of controlling. As shown in table 2, at the significance level of 0,05, the null hypothesis in favour of the alternative was rejected based on a p-value very close to zero and. Therefore, we can argue that companies with implementation of controlling have a higher number of employees (see the first line in the table). The table 2 also shows that the size of enterprises does not affect implementation of controlling in all areas of business. For example, in sectors such as transport and storage; information and communication activities, services, construction industry and agriculture, forestry, fishing, the null hypothesis has not been rejected, as the p-value is always above the significance level. On the other hand, in sectors such as administrative and support activities; professional, scientific and technical activities; wholesale and retail and manufacturing, H:0 has been rejected in favour of an alternative hypothesis, and we can say that the size of a business in these divisions of enterprises varies with respect to the impact of implementation of controlling.

Table 2: Mann-Whitney U test – controlling in selected divisions of enterprises

Division of enterprise	Contr. YES	Contr. NO	U	Z	P-value	Z edited	P-value edited
All divisions of enterprises	323	269	30448,0	-6,2484	0,0000	-6,2515	0,0000
Administrative and support activities	12	14	43,0	-2,0831	0,0372	-2,0849	0,0371
Transport and storage	13	12	43,5	1,8494	0,0644	1,8501	0,0643
Information and communication activities	30	21	251,5	1,2058	0,2279	1,2071	0,2274
Professional, scientific and technical activities	33	38	430,0	2,2654	0,0235	2,2687	0,0233
Services	13	16	81,5	0,9648	0,3347	0,9660	0,3341
Construction industry	39	26	406,0	-1,3458	0,1784	-1,3472	0,1779
Wholesale and retail	60	46	985,5	2,5115	0,0120	2,5153	0,0119
Agriculture, forestry, fishing	7	21	65,5	-0,3979	0,6907	-0,3984	0,6903
Manufacturing	116	75	2788,5	4,1839	0,0000	4,1851	0,0000

Source: Own research

The graphical representation in the form of box plots (figure 1) also visually confirms this claim. Those charts show the vertical shift of median upwards for enterprises with implementation of controlling (in the divisions of enterprises Manufacturing and Wholesale and retail). On the other hand, the service industry did not manage to statistically demonstrate the difference in both groups, but from the fourth box plot, it can be concluded the median and upper quartile displacement is higher compared to enterprises without implementation of controlling. Completely different results are in division of enterprise Agriculture. This industry did not manage to statistically demonstrate the difference in both groups, but from the third box plot, it can be concluded the median and upper quartile displacement is lower compared to enterprises without implementation of controlling.
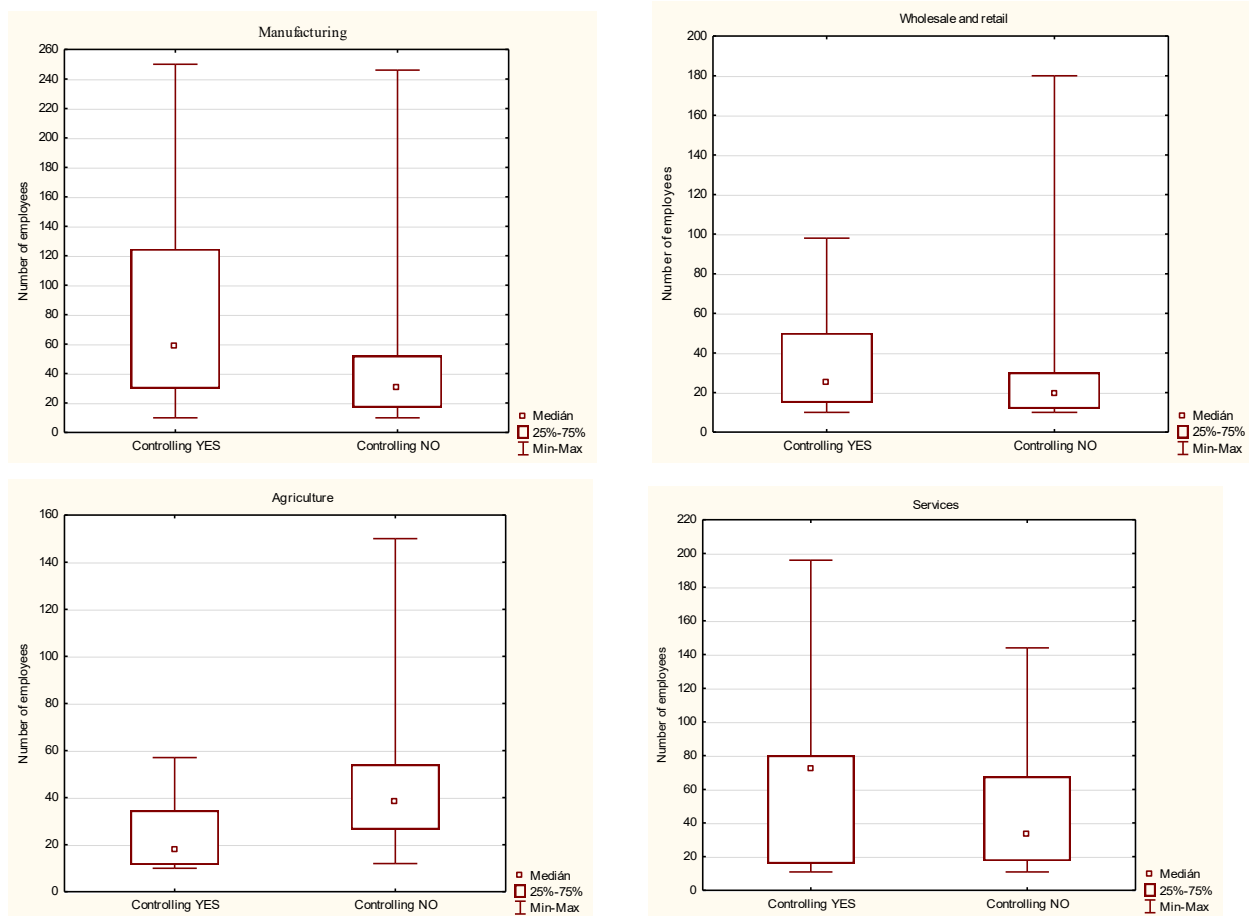
Figure 1: Mann-Whitney U test – controlling in selected divisions of enterprises (upper - Manufacturing and Wholesale and retail, downer - Agriculture and Services); Source: Own research

With the research in the area of enterprise size, many authors have dealt with it (Stock, & Seliger, 2016; Frazzon, Hartmann, Makuschewitz, & Scholz-Reiter, 2013) and with the research in the fiels of the business sector as well (Muller, 2019; Prause, 2015), who says that this factors influence the enterprises. Similarly, the results of Bui, Tu Le, & Nguyen (2020) have shown that 100% of Vietnamese businesses use management accounting, which is part of the financial controlling. They found that the level and content of managerial accounting applications in businesses varied by size of the company and by the business divisions. In practice, this research can help small and medium-sized enterprises in particular if they decide to implement controlling or not.

5. Conclusion

This paper primarily deals with the analysis of the relationship between enterprise size and implementation of controlling and the analysis of the relationship between the divisions of enterprises. When evaluating the size of the enterprise, statistically significant difference has been found between the enterprises with implementation of controlling and without implementation. Most of enterprises with implementation of controlling employ more workers. Other practical findings related to the divisions of enterprises into nine groups according to the business sector. It was found a significant difference between the size of the company and the implementation of controlling in divisions of enterprises such as administrative and support activities; professional, scientific and technical activities; wholesale and retail and manufacturing. Only the agriculture sector showed a completely different trend. Based on the graphic expression it has been found that in this sector

(although this was not statistically confirmed at the significance level $\alpha = 0.05$), the higher number of employees does not mean controlling implementation.

6. References

Budíková, M., Králová, M., & Maroš, B. (2010). Průvodce základními statistickými metodami. 1st ed. Praha: Grada Publishing.Callahan, R. (1962). Education and the Cult of Efficiency. Chicago: University of Chicago Press.

Bui, N. T., Tu Le, O. T., & Nguyen, P. T. T. (2020). Management accounting practices among Vietnamese small and medium enterprises. Asian Economic and Financial Review, 10(1), 94–115. https://doi.org/10.18488/journal.aefr.2020.101.94.115.

Cornel, G., & Lavinia-Maria, G. (2012). Controlling - A useful tool for top management. Analele Universitatii Maritime Constanta, 13(18), pp. 305-308.

Čebukina, M. V., Shatalová, T. N., Jirnová, T. V., & Bobkova, E. Y. (2013). Controlling as a tool for implementation of the system for the enterprise resource potential management in its capitalized form. World applied sciences journal, 27 (4). 444-447.

Davila, A., & Foster, G. (2005). Management accounting systems adoption decisions: Evidence and performance implications from early-stage/startup companies. Accounting Review, 80(4), 1039–1068. https://doi.org/10.2308/accr.2005.80.4.1039.

Drury, C. (2015). Management and cost accounting. 9th Ed. Andover: Cengage Learning.

Fibírová, J. (2003). Reporting. Moderní metoda hodnocení výkonnosti uvnitř firmy. Praha: GRADA Publishing a.s.

Frazzon, E. M., Hartmann, J., Makuschewitz, T., & Scholz-Reiter, B. (2013). Towards Socio-Cyber-Physical Systems in Production Networks. In 46th CIRP Conference on Manufacturing Systems 2013, 7 (0), 49–54.

Freund, R. J., Wilson, W. J., & Mohr, D. L. (2010). Statistical methods. 3rd ed. Amsterdam: Elsevier.Garrison, R., Noreen, E., & Brewer, P. (2012). Managerial accounting. 14th ed. New York: McGraw-Hill/Irwin.

Horváth, P. (2003). Controlling. München: Vahlen.

Král, B. & kol. (2010). Manažerské účetnictví. Praha: Management Press.

López, O. L., & Hiebl, M. R. W. (2015). Management accounting in small and medium-sized enterprises: Current knowledge and avenues for further research. Journal of Management Accounting Research, 27(1), 81–119. https://doi.org/10.2308/jmar-50915.

Magretta, J. (2002). Why business models matter a conversation with robertredford. Harvard Business Review, 58, 86–92.

Müller, J. M. (2019). Business model innovation in small- and medium-sized enterprises: Strategies for industry 4.0 providers and users. Journal of Manufacturing Technology Management, 30(8), 1127–1142. DOI: 10.1108/JMTM-01-2018-0008.

Prause, G. (2015). Sustainable business models and structures for industry 4.0. Journal of Security and Sustainability Issues, 5(2), 159–169. DOI: 10.9770/jssi.2015.5.2(3).

Reichmann, T. (1997). Controlling: Concepts of Management Control, Controllership, and Ratio. Heidelberg: Springer-Verlag Berlin.

Stock, T., & Seliger, G. (2016). Opportunities of Sustainable Manufacturing in Industry 4.0. Procedia CIRP, 40, 536–541. https://doi.org/10.1016/j.procir.2016.01.129.

Simkin, L. (2010). People and processes in marketing planning: The benefits of controlling implementation. Journal of Marketing Management. https://www.tandfonline.com/doi/pdf/10.1080/0267257X.1996.9964423?needAccess=true.

Stoner, A. F. J., Freeman, R. E., & Gilbert, D. R. Jr. (1998). Management. India: Prentice-Hall of India Pvt. Ltd.

Vlčková, M. (2018). Which companies have implemented management accounting? IDIMT-2018 - Strategic Modeling in Management, Economy and Society, 26th Interdisciplinary Information Management Talks, Sept. 5–7, 2018, Kutná Hora, Czech Republic, (pp. 417 - 425). ISBN 978-3-99062-339-8.

Wirtz, B. W., Pistoia, A., Ullrich, S., & Göttel, V. (2016). Business Models: Origin, Development and Future Research Perspectives. Long Range Planning, 49(1), 36–54. https://doi.org/10.1016/j.lrp.2015.04.001.

SUSTAINABILITY REPORTING AND MANAGEMENT OF GERMAN AND THE BALTIC RETAIL CHAINS: WHAT ARE THE DRIVERS?

Lina Dagilienė, Viktorija Varaniūtė

Kaunas University of Technology, Lithuania
lina.dagiliene@ktu.lt; viktorija.varaniute@ktu.lt

Judith Pütter

Hochschule Hamm-Lippstadt, Germany
judith.puetter@hshl.de

Toomas Haldma

University of Tartu, Estonia
toomas.haldma@ut.ee

Keywords

Drivers, sustainability reporting, sustainability management, retail chains

Abstract

Retailers exert a significant influence on sustainability issues due to their position in the supply chain between producers and consumers. This paper aims to explore drivers influencing sustainability management and reporting in retailing chains by using the neo-institutional theory approach. The data collected through desk research and semi-structured interviews with retail chains in the Baltic countries and Germany are expected to provide insights into drivers motivating retailers to implement sustainability practices and report on them. Overall, the key drivers are environmentalism-related (energy efficiency, logistics and packing); product-related (healthy products, origin of products); followed by drivers related to consumer responsibility and behavior. The results slightly differ across the countries examined, and therefore different implications for retails chains could be offered.

1. Introduction

Due to the position in the supply chain between producers and customers, retailers have a high interest in safeguarding sustainable behavior (Wilson, 2015). The increasing policy traction of the Sustainable Development Goals (SDGs), and especially SDG 12 on responsible consumption and production, foster the proliferation of environmental sustainability strategies in the retailing sector. Companies will be expected to report on their sustainability progress (Naidoo & Gasparatos, 2018). To our knowledge, only a few papers address the issues of sustainability reporting (*further SR*) in retailing sector (Baker et al., 2012). A majority of the research explores information available on the

Internet and publicly available reports from the world's largest retailers (Björklund et al., 2016) or retails chains (Utgård, 2018), mostly in large developed countries (Filimonau & Gherbin, 2017). We can easily ascertain how organizations present themselves on the Internet, because this forum allows them to communicate their purposes and their way of operating (Utgård, 2018), but there is little systematic knowledge about the drivers that influence retailers to report on sustainability issues publicly.

In this study, we concentrate on SR of retail chains that operate in the Baltic countries and Germany by using a comparison approach. Naidoo and Gasparatos (2018) emphasize the lack of national/regional case studies, because the current studies of environmental reporting tend to privilege Anglo-American countries (Mata et al., 2018) and other Western European countries with historically higher level of social responsibility, democracy, and market economy in comparison to Eastern European countries. Moreover, there is a lack of qualitative studies that explore causal consequences in retailers' activities and explain the drivers for adopting sustainability practices and reporting on them. To resolve this gap, we raise a research questions in particular:

RQ1 What and how do the drivers influence sustainability management and reporting practices of retail chains in the Baltic countries and Germany?

Firstly, we contribute by exploring the state of the art of SR practices of retail chains across several countries. Secondly, we explain how different general and sector-specific drivers are expressed (in each country) and how they influence SR of retail chains. We focus on sector-specific SR by observing the views of retail chains (interviews), supported by publicly available documents.

2. Theoretical background

Previous research on SR in retailing. When it comes to the retail sector, there is a considerable amount of research dedicated to retailers' sustainability in their supply chains (Björklund et al., 2016; Wilson, 2015): logistics-related environmental considerations and environmental performance indicators (Björklund et al., 2016); energy-saving responsibility in retail environments; and food waste management practices (Filimonau & Gherbin, 2017).

Most of the research is dedicated to the content analysis of web information and sustainability reports of the largest retailers. By using the institutional theory approach, Baker et al (2012) analyze sustainability reports produced by large retail distribution companies. Björklund et al. (2016) argue a need for in-depth case studies of retailers to understand corporate environmental consciousness to investigate the dialogue between the information creators of sustainability reports and those who are responsible for their content. Utgård's (2018) study reveals that foreign chains, chains using private brands and vertically integrated chains are more likely to communicate sustainability information on their web pages. In addition, Filimonau and Gherbin (2017) employed content analysis of corporate materials and a qualitative method of primary data collection to explore managerial attitudes and approaches to food waste mitigation in supermarkets.

Drivers influencing the implementation of SR in retailing. The institutional theory focuses on the analysis of external factors of social and economic environment and their impact on the company in order to understand the relations between the behavior of the company and the broader social environment in which it operates, as well as the company's impact on the environment (Higgins & Larrinaga, 2014). Based on the review of the scholarship, we can assume that there are drivers that affect SR both in *general* and in *sector-specific* ways.

Odera et al. (2016) investigate factors influencing corporate SR in general. The key factors were corporate image making and avoiding government scrutiny, which could lead to adverse publicity. More significantly, Odera et al. (2016) note that SR is considerably diverse among companies in

developed countries, compared to developing and transition countries. Another comprehensive theoretical review, conducted by Hahn and Kühnen (2013), systemized various external and internal determinants that influence SR. It is important to notice that the company's sector affiliation is the most frequently addressed external determinant (Hahn & Kühnen, 2013).

On the other hand, some drivers are more relevant to the *retail sector*. Jones and Comfort (2018) emphasize the importance of storytelling in addressing the visibility of retailers', in order to create a strong relationship with the company's stakeholders. By investigating sustainability reports of leading US retailers, Jones and Comfort (2018) review the way stories have been used in sustainability reports, and they identify a number of storytelling formats – namely, photographs, cameo case studies and impact stories. Thus, we assume that retail companies see SR as an institutionalized way to tell a story to their stakeholders or to create a better corporate image.

The supply chain position is another important driver for SR in retailing. Hahn & Kühnen (2013) admit that research on the position of the value chain as a driver is still scarce. Again, retailing companies may have an additional motivation to behave in a sustainable way and to perform SR. Utgård (2018), based on a sample of 208 Norwegian retail chains, emphasize that foreign ownership, being a private brand and the chain's structure influence whether the company engages in SR on their web pages.

Overall, expectations from internal and external stakeholders become a dominant driver in retailing. From an external point of view, environmental organizations, consumers, the media, and others put pressure on retailers to increase accountability and transparency. This is especially the case in reducing the sustainability impacts of retailing operations through the implementation of environmental management systems and environmentalism-related measures and indicators. For example, energy use represents one of the highest operational expenses for retailers (Naidoo & Gasparatos, 2018). The need to reduce energy costs is also consistent with the motive of financial gains, mainly due to cost savings incurred from reducing operational expenses. Retailers also seek to implement waste reduction strategies, with the aim of reducing waste disposal costs, as they must appropriately dispose of substantial amounts of plastic, cardboard and food waste (Naidoo & Gasparatos, 2018). According to Ruiz-Molina and Gil-Saura (2018), Spanish retailers (with the top sectors being grocery and apparel) are making efforts to design and implement sustainable practices from a triple bottom line framework.

It is important to note that – directly or indirectly – most of the aforementioned drivers revolve around expected financial gains. However, significant cost savings and increased profitability may only occur after substantial initial investments are made to change the current systems. According to Naidoo & Gasparatos (2018) regulatory pressures seem to drive sustainability implementation up to a certain extent, although this does not explain why some companies decide to go beyond measures required by law.

3. Research methodology

The main objective of this research is to explore drivers influencing retail chains' sustainability management and reporting implementation. Therefore, the nature of this research is explanatory. Our research strategy is focused on the application of *qualitative methods by analyzing cases studies* of the retail chains in Baltic countries and Germany. The research is divided into two parts: desk research analysis and interview analysis. Both for the desk research analysis and for interview analysis, we carried out the main stages for qualitative data analysis – namely, collecting data, open coding, axial coding and developing insights.

Data were collected and analyzed in 2019. For desk research we analyzed the information (only at the country level) provided on retail chains' websites (10), annual reports (8) and CSR report (1). What is more, a semi-structured interview was used to reveal the drivers influencing retail chains' sustainability management and reporting (*further SMR*). The interviews were conducted in the local language. In total, 13 interviews were conducted (3 of respondents work in a position directly related to social responsibility or sustainability; 4 respondents work in strategic-level positions; other respondents came from a variety of fields). The total duration of interviews lasted for 849 minutes. All the interviews were recorded and transcribed. The total number of transcript pages of interviews were 146,50. After the transcription of all the interviews was completed, the transcripts were analyzed by researchers separately via a systematic process of coding and categorization intended to group the information from the transcripts into similar concepts or themes that emerged from the analysis. We then identified keywords, which were prevalent in all interviews using the MAXQDA program (or manually).

4. Research results

Main findings of desk research analysis. The results contain short an overview of legal requirements for SR, how sustainability is expressed in public strategies of the selected retail chains and, finally, drivers for sustainability management (*further SM*).

Legal requirements. An overview of legal requirements regarding reporting non-financial/ sustainability information in Baltic countries and Germany revealed minor differences amongst the countries. By analyzing the legal requirements regarding the reporting of non-financial/ sustainability information in Baltic countries and Germany, we observed that the countries' requirements vary – from requirements to publish stand-alone reports on sustainability (Germany) to incorporating sustainability in other compulsory reports (Baltic countries). The requirements present in Baltic countries and Germany are generally set out for large companies with more than 500 employees and public interest entities.

Strategy. In order to determine how sustainability is embedded in the retail chains' strategy, a range of sources was analyzed: CSR, management reports, homepages and media/press releases. The source was selected based on availability. The priority was on the CSR and management reports, and if they were unavailable, webpages were analyzed. Sustainability information in the strategies of each country's retail chains is analyzed and key elements are summarized in Table 1.

In summary, Baltic countries and Germany retail chains publicly declare that they aspire and strive to be sustainable by emphasizing this in the aims and strategy of their companies. However, it is important to notice whether retail chains in Baltic countries and Germany provide information of sustainability in reports.

Table 1. Summary of key elements of sustainability information in strategies of retail chains in Baltic countries and Germany

Country	Key elements of sustainability in strategies
Estonia	environmental sustainability; clients; innovative solutions; socially responsible business (social responsibility); social and environmental aspects of management; relations with stakeholders;
Latvia	closest store; serve our clients' needs; backyard shop; healthy; favorable and safe environment. secure welfare; help for families;
Lithuania	responsible (or socially responsible); taking care of environment; being a part of society; favorable work conditions for employees; long-term partnerships; local produces; healthier,

	organic produces: the supply chain to meet environmental standards; resource conservation;
Germany	organic products; support the farmers in our region; sustainability; responsibility: environmentally friendly; reducing our ecological footprint.

Sustainability management. Baltic countries and Germany retail chains publicly declare that they aspire and strive to be sustainable by emphasizing this in the aims and strategy of their companies. However, analysis of different reports has shown that retail chains in Baltic countries and Germany comply the law by providing a minimum information of sustainability in reports. Desk research results lead to the importance of SR of retail chains that to ensure not only transparency but also the dissemination of information on sustainability to different stakeholders. Due to this, it is important to get the deeper insights from interviews with retail chains that to know how to foster the SR.

Main findings of interview analysis. The results of the interviews contain potential drivers of SR, as well as the drivers for SM identified based on retail chains in Baltic countries and Germany.

Potential drivers of SR. Analysis of the interviews revealed that SR can be expressed by three main potential drivers: regulations, enabling and implementation). 'Regulations' refers to the information that is relayed to the national regulative bodies and the legal acts that influence sustainability (such as recommendations, directives, sustainability policies, surveys, laws, requirements and procedures, etc.). 'Enabling' refers to the information that is related to enabling sustainability solutions, such as the strategies, communications, advertising, training, and workplace culture, etc. 'Implementation' refers to the information that is related to the providing results about the company in the report submitted to management, external stakeholders, and others via accounting systems, reports and sustainable solutions, etc. From desk research we can see clearly that sustainability elements are included into strategy (publicly), but still retail chains do not issue separate sustainability reports, only fragmental communication through media channels.

When analyzing the interview data (Figure 1), it can be observed that the main potential drivers of SR are as follows: regulations (41.14%); implementation (32.73%); and enabling (26.13%). This means that for the retail chains analyzed in this study, it is important not only to meet SR requirements, but also to enable SR via different solutions, as well as to implement it by communicating results to others.

The analysis of Estonian retail interviews shows that the focus is on the regulations (45.70%; EST 1: "<…Our parent company is a listed company and prepares a quarterly report for the stock Exchange. Part of this report is a sustainability report and each country gives its input quarterly...>") and implementation (42.20%; (EST 3: "<… we do have projects to develop sustainability, but we don't contribute to the SM…>"). The least attention is paid to the enabling (12.10%) of SR (EST 3: "<...We use inspirational video clips on the web on how to consume and buy economically>"). The lower attention on the enabling role of SR can be explained by the fact that only one company had the sustainability specialist position, who is responsible for the sustainability report compilation, and had more systematic view on the sustainability as enabler.
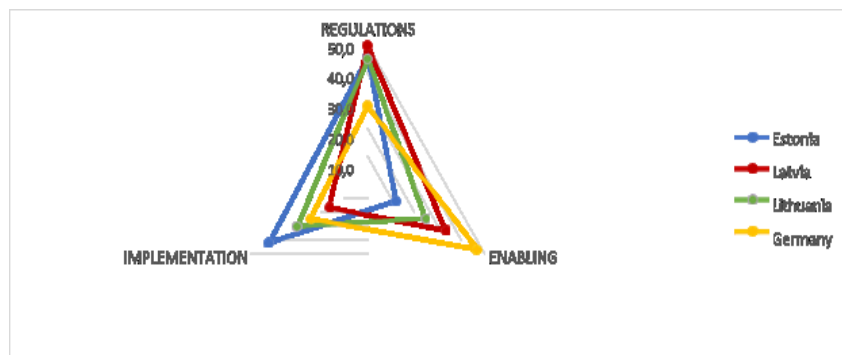
Figure 1. Potential drivers of SR in Baltic countries and Germany

By analyzing Latvian retail chains interviews, we can notice that the main focus is on the regulations (50%) which is policy of one retail chain (LV 2: *"<...sustainability report will be created within the UN benchmark framework...>"*) and enabling (33,33%; LV 3: *"<...employee wellbeing, youth support through various initiatives, employee and customer education etc. ...>"*). The least attention by Latvian retail chains is paid on the implementation (16.67%) of SR.

The interview analysis revealed that Lithuanian retail chains focus on regulations (45.24%) of SR (LT 1: "<... In purchasing decisions (products, energy, transport, etc.), it is required to take into account sustainability policy recommendations in the relevant areas ...>"; LT 2: "<... we have a very clear procedure, in the sense that we even have a guidebook and how, who and why should sort ...>"). Although every potential drivers of SR are important, Lithuanian retail chains paid the least attention to the enabling (24.60%) of SR (LT 4: "<...posters, communication are all launched, salespeople are trained that this is a day of ecology, sustainability...>").

The analysis of German retail interviews revealed that the main focus is on the enabling of SR (47.06%; GER 1: "<...it is important that people who work for me are especially trained, it is necessary that the stand behind "Bio" to explain what lies behind sustainability...>"). German retail chains paid the least attention to the implementation of SR (24.71%; GER 2: "<...I cannot say that we have a complete and structured process [of reporting] It would be an excessive demand...>") and regulations (28.24%; GER 3: "<...there would have been much more pressure, if the EU CSR Reporting guidelines had been valid for us as well...>"). Regulations as well as the implementation of reporting system seem to be less relevant for the regarded retail chains, as they were either too small or due to their organization form are not addressed by SR regulations.

As SR is voluntary for retail chains, retail chains do not issue stand-alone reports and properly integrate into annual / management reports. They do fragmental reporting duties, derived from the law (for example, report on waste, report on package for tax authorities etc.).

Drivers for the sustainability management. Analysis of the drivers of SM in all selected countries (Figure 2) revealed that the main drivers motivating retail chains to be more sustainable are energy efficiency and logistics and packing (18.20%); healthy products (17.80%); origin of the products (12.74%); and consumer behavior and responsibility (12.40%).
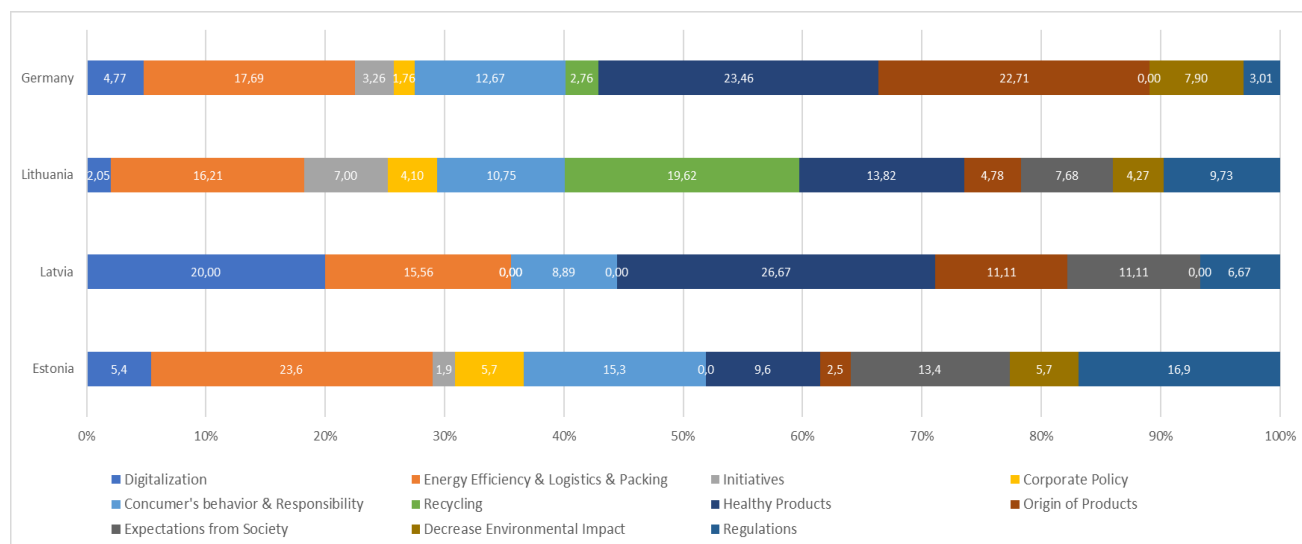
Figure 2. Drivers of SMR in Baltic countries and Germany

Analysis of the interviews revealed that retail chains in Estonia emphasize the following drivers as most important: energy efficiency, logistics and packing (23.60%); regulations (16.90%); consumer behavior and responsibility (15.30%). These three drivers are very important due to the competitive market, where high attention on the efficiency are important; recent changes in the regulations (such as packaging) and general consumers trends (as explained by the one of the interviewees "<…there is no backwards moving that we or our customers are less eco-friendly…>").

Retail chains in Latvia emphasize the following drivers as most important to SMR: healthy products (26.67%; LV 1: "<…special shelf in the store for bio, eco and special products>"); digitalization (20.00%; LV 2:"<…possibility to pay for different bills (electricity, utilities, telecommunications etc.)…>"); and energy efficiency, logistics and packing (15.56%; LV 3: "<… emphasis is placed on sorting (cardboard, plastic, household waste) actively sorting waste, buying service from the waste manager…>"). The analysis of Latvian retail chains shows that the main drivers are related to the environmental-legal factors and digitalization.

Lithuanian retail chains identify the following drivers as most important: recycling (19.62%); energy efficiency, logistics and packing (16.21%); and healthy products (13.82%). These drivers are very significant because they lead to the fulfilment of requirements for SR. Recycling as a driver is expressed by sorting (LT 1: "<...the staff are familiar with the principles of sorting and all packaging waste....>"); bio-waste (LT 3: "<...part of the waste is handed over to a food waste management company for proper disposal…>"). Energy efficiency, logistics & packing is expressed by energy saving (LT 2: "<...we have different temperature zones, no door left open, no cracks...>"), packing (LT 1: "<...we already offer disposable tableware alternatives from paper, sugar cane to plastic buyers...>"), and logistics by using different types of vehicles (LT 2: "<...when purchasing energy, priority shall be given to the purchase of green energy, or only hybrid cars...>"). Healthy products as a driver is expressed by the sustainable products (LT 3: "<...is such a healthy product line, it is expensive enough...>"), sugar free (LT 4: "<...where sugar is less, well that keyhole thing...>"). Retail chains in Lithuania do not place as much of an emphasis on the supply chain factor (origin of products).

Retail chains in Germany emphasize the following drivers as most important to SMR: healthy products (23.46%); origin of products (22.71%); and energy efficiency and logistics and packing (17.69%). Most of the "organic retail chains" are historically justified by the demand of "organic food" (GER 1: "<...the development of organic products was quite fast...>"). It is especially important for the owner of retail chain GER 1 to sell fresh food (GER 1: "<…I love to sell good/fresh food…>").

The business model of retail chain GER 1 and GER 2 is based on 100% "organic/bio food" (GER 1: "<…we carry a 100% organic product range […] all products are certified...>"). Also, a new listing of a vegetarian product was very successful: (GER 3 "<...we had this vegetarian sausage, which hit the market like a bomb...>"). The origin of the product is somehow accompanying the "healthy product". Especially for organic retail chains are the origin of cultivation or the farmer an important factor to guarantee a healthy product and to provide transparency (GER 1: "<…is important for me, that distinguishes me from my competitors. They cannot say that, or they cannot order from a regional farmer…>"). Finally, energy efficiency, logistics and packing have also a great relevance. The development of new packing systems should lead to waste reduction (GER3: "<...we use reusable packaging, in order to reduce packaging waste...>"). On the other hand retail chains need to comply with the legal packaging laws (GER 3:"<...due to the packaging law a new process was initiated to develop recyclable packaging...>").

Interview research results show that both Baltic countries and Germany have many initiatives related to SM, which are influenced by main drivers motivating retail chains to be more sustainable are energy efficiency and logistics and packing, healthy products, origin of the products and consumer behaviour and responsibility. On the other hand, these initiatives are declared on the organizations' websites, however retail chains do not provide any official statistics and information on sustainability in their annual / management reports.

5. Discussion and conclusion

Our results show that concerns about environmental impact and so-called environmentalism- related factors – energy efficiency, packing and recycling – are prevalent across retail chains in all the investigated countries. The results are in line with previous studies (Naidoo & Gasparatos, 2018, Filimonau & Gherbin, 2017). For example, energy use represents one of the highest operational expenses for retailers (Naidoo & Gasparatos, 2018). Therefore, many retailers focus on energy conservation and greenhouse gas emission reductions. Moreover, packaging and food waste are becoming priority sustainability issues for retailers, as well. The need to reduce energy costs is also consistent to the motive of financial gains, mainly due to cost savings from reducing operational expenses.

The other important group of drivers includes "healthy products" and "origin of products". Again, this is a result of a mixed impact from regulatory pressures (for example, to limit quantity of sugar) and expectations from consumers, as well as other actors across the supply chain (e.g. origin of the product). Retail chains seek to create stronger relationship with customers (Jones & Comfort (2018), Utgård (2018)) and are more consumers orientated rather than producers orientated. Therefore, they react / adopt sustainability initiatives by changing products' assortment towards healthier product, pay more attention to the origin of the product.

Retail chains acknowledge the regulatory risks, as the chain is present in several countries, and emphasizes safety regulations, such as product traceability, third-party certification and laboratory tests. Compulsory regulation and normative practices (Naidoo & Gasparatos, 2018; Ruiz-Molina & Gil-Saura, 2018) were confirmed to be important drivers. As only the largest companies must consistently provide SR within legal requirements, smaller retail chains therefore do not issue standalone sustainability reports. In addition, interview data reveal the main potential drivers of SR such as regulations, implementation and enabling. Our results show that drivers affecting SM are mostly normative and consumer-driven, while potential drivers that would foster SR are regulatory.

Further research. Having chosen a qualitative approach prevents a broader data collection method, which may provide different views. It would be worthwhile to carry out further empirical analyses of sustainability accounting and reporting through a detailed (longitudinal) case study of retail chain

or a quantitative survey to gather a broader range of insights. The interviewed representatives from retail chains confirmed the inclusion of sustainability data into their internal accounting systems, while most of them do not seek to report on these data publicly.

6. Acknowledgments

7. References

Baker, C.R., Cohanier, B., and Pederzoli, D. (2012). Corporate Social and Environmental Reporting in the large retail distribution sector. Procedia Economics and Finance, 2, 209-218.

Björklund, M., Forslund, H., and Isaksson, M.P. (2016). Exploring logistics-related environmental sustainability in large retailers. International Journal of Retail & Distribution Management, 44 (1), 38-57, https://doi.org/10.1108/IJRDM-05-2015-0071

Filimonau, V., and Gherbin, A. (2017). An exploratory study of food waste management practices in the UK grocery retail sector. Journal of Cleaner Production, 167, 1184-1194.

Hahn, R. and Kühnen, M. (2013). Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. Journal of Cleaner Production, 59, 5-21.

Higgins, C. and Larrinaga, C. (2014). Sustainability reporting: insights from institutional theory, in Sustainability accounting and accountability. Routledge, New York, N.Y., 273-285.

Jones, P. and Comfort, D. (2018). Storytelling and sustainability reporting: an exploratory study of leading US retailers. Athens Journal of Business and Economics, 4 (2), 147-162.

Mata, C., Fialho, A., and Eugenio, T. (2018). A decade of environmental accounting reporting: What we know? Journal of Cleaner Production, 198, 1198-1209.

Naidoo, M. and Gasparatos, A. (2018). Corporate Environmental Sustainability in the retail sector: Drivers, strategies and performance measurement. Journal of Cleaner Production. In Press. https://doi.org/10.1016/j.jclepro.2018.08.253

Odera, O., Scott, A. H., and Gow, J. (2016). Factors influencing corporate social and environmental disclosures: a systematic review. International Journal of Business Governance and Ethics, 11, 116-134.

Ruiz-Molina, M. and Gil-Saura, I. (2018). Sustainable Practices in Spanish Retailing: Evidence from Apparel and Grocery Retailers. Lavorata, L. and Sparks, L. (Ed.) Food Retailing and Sustainable Development, Emerald Publishing Limited, 23-43. https://doi.org/10.1108/978-1-78714-553- 520181002

Utgård, J. (2018). Retail Chains' Corporate Social Responsibility Communication. Journal of Business Ethics, 147 (2), 385-400. https://doi.org/10.1007/s10551-015-2952-2

Wilson, J.P (2015). The triple bottom line: Undertaking an economic, social, and environmental retail sustainability strategy. International Journal of Retail & Distribution Management, 43(4/5), 432-447. https://doi.org/10.1108/IJRDM-11-2013-0210.

INTEGRATION OF STRUCTURAL PRINCIPLES OF RESPONSIBILITY MANAGEMENT INTO LIQUIDITY MANAGEMENT

Lucie Brabcová

Faculty of Finance and Accounting
University of Economics, Prague
lucie.brabcova@vse.cz

Keywords

Abstract

Ongoing discussions are held among academics and professionals on the required degree of centralization of liquidity management activities. The purpose of this contribution is to offer another view into this discussion by adding an interdisciplinary dimension of responsibility management as a determinant of liquidity management centralization.

Traditional responsibility centers are transformed into liquidity centers (inflow center, outflow center and self-financing center on operational side, and low-scope investment center, high-scope investment center and "not an investment center" on the investment side) and presented in a matrix demonstrating the possibilities of the company structure and level of liquidity management centralization by means of decentralized financial investment function. An issue of holding precautionary cash balances leading into losing the benefit of economies of scale for investment purposes is addressed as well while summarizing the need for liquidity centralization by element.

1. Introduction

A degree of liquidity management centralization is one of the key topics of corporate financial management, frequently argued by academics and professionals in the financial management field. The purpose of this contribution is to offer another view into the discussion by adding an interdisciplinary dimension. This is to be achieved by using a building block of responsibility management, economic structure, defined and represented by responsibility centers. The main point of our interest are business combinations with international complex structures, operating in multiple markets. Their subsidiaries are usually acting as separate legal entities, however, their independence in this area of management is often compromised as they have to comply primarily with the ultimate goals of the group (Roun, 2007).

1.1. Methodology

To achieve the aim of this contribution, an analysis of literature was performed, followed by a case study. The analysis of literature sources focuses on business' organizational and economic structures. After that, the recognized principles are applied to two different company structures with an emphasis

on merging the approaches from management accounting and liquidity management disciplines. The conclusions of the paper are then based on the author´s experience from management of the business combinations, as well as induction of general principles from the described cases. The research question we are striving to answer is: How can economic structure, represented by responsibility centers, be translated into centralization of corporate liquidity management?

The key part of the literature review, i.e. centralization of liquidity management, is based on search in Scopus and Web of Science. Keywords used are "treasury" or "liquidity" or "cash management" combined with "centralisation" or "centralization", i.e. "treasury AND centralisation". These six combinations are searched as a topic (WoS) or title-abs-key (Scopus). Further limitation of categories (economics and business finance) is made, while no limitation of publication years is used. This search results in total of 32 papers (Scopus: 24 results, WoS: 8 results). Also, papers published in Journal of Payments Strategy & Systems and Journal of Corporate Treasury Management were scanned. Because of their narrow specialization, the search was limited to "centralisation OR centralization", resulting in nine papers. Due to the results overlap and the topic specifics, not all the papers have been selected to the literature review. Additional sources of literature are used as well.

To apply the principles of responsibility centers into liquidity management, let us presume two extreme variants of legal entities in an international group structure, where both examples are based on generalized and anonymized structures of real companies. In Group A the subsidiaries act as separate, individually operating businesses, generating both inflows and outflows from their operations. They might have intercompany relationships resulting in cash flows, but these are not of key volume. In Group B the entities are split into sale companies generating large inflows and irrelevant volume of outflows related to operating expenses, and production companies with large outflows funding their operations. The intercompany cash relationships are key for Group B.

In the following text, terms cash management, liquidity management and treasury management are interchangeable, unless specified differently. Use of netting and cash pooling tools is assumed.

2. Literature review

2.1. Organizational structure

According to Král (2018), a decision on organizational and economic structure of the company belongs among prerequisites of building an efficient system of responsibility management. The organizational structure is defined by vertical and horizontal structures. Horizontal structures determine coordination and division of the activities within the company. Two types of horizontal structures are identified: functional and divisional. Divisional structure consists of primarily autonomous departments that are able to manage a region or the full product process mostly independently. Departments of this structure are typically managed as decentralized. In functional structure, the subsidiaries are usually parts of a coordinated functional chain with strong horizontal cooperation and need for centralized coordination (ACCA & Kaplan Publishing UK, 2019; Blažek et al., 2010; Král, 2018). Šoljaková (2009) recognized a trend of leaning from functional structures to divisional structure. Vertical structures, that are key to responsibility distribution, are influenced by the level of centralization (Král, 2018). Blažek (2010) identifies general factors influencing the degree of centralization: intensity of flows of products, capital, finances and know-how. Christie et al. (2003) found evidence that "knowledge transfer costs being, on average, relatively more important than control costs in the decentralization decision" (p. 33; highlighting removed by the author). Jensen & Meckling (1995) introduced the dimension of control costs and knowledge transfer costs.

2.1.1. Centralization of liquidity management

A continuous trend of centralizing liquidity management has been identified by multiple authors (Polak, 2009b, 2010; van Alphen & Huiskes, 2008; Venkat & Baird, 2016; Wessels & Van Rooyen, 2010; Westerman & von Eije, 2002). This long-term trend has been indicated already by Soenen (1986), being in line with his references from earlier years (Gitman et al., 1979; Reed, 1972; Smith & Sell, 1980). Among external factors accelerating this trend can be considered euro zone establishment (Westerman & von Eije, 2002), harmonization of SEPA (Barbas, 2009; van Alphen & Huiskes, 2008) and developed technologies (Milkau & Bott, 2015; Regino & Bernardelli, 2008; van Alphen & Huiskes, 2008), giving rise to new banking products, such as payment factories (Cobben, 2008). In the below section are presented opinions on the liquidity management centralization question, along with related advantages and disadvantages.

Brealey (1992) states that "good cash management requires centralization to some extent" (p. 821; translated by the author), especially in the question of cash monitoring and maintenance of bank relationships. As Polak (2009b) summarized, "cash management is an activity that clearly benefits from economies of scale" (p. 109). Centralization is also supported by Wagner (2009), who mentions that it is usually more rational to manage liquidity on a centralized level, referring to a situation when one subsidiary generates excess cash, while another needs funding via short-term loan. Adalsteinsson (2014) comments that there is no unified approach relevant and appropriate for all. He believes that a certain level of centralization is necessary not only to support malfunctioning subsidiaries, but also to maintain a shared approach. This means that centralization shall be applied at least to risk appetite, group guidelines, standards and monitoring. The decentralized approach is then appropriate for entities that "are self-funded, or attract most of their funding from the local markets" (Adalsteinsson, 2014, p. 15). Soprano (2015) describes these as "business cases where some companies are structurally in liquidity excess or structurally in funding need" (p. 152). Van Alphen & Huiskes (2008) share results of a market survey, where the following drivers for centralization of liquidity management were identified: Use of excess cash for internal funding, cash control, cash visibility, enhance yields/reduce interest costs (>50 % of the respondents), increased efficiency, optimise balance sheet (>30 %), reduce account complexity, tax/regulatory issues and company culture. Adalsteinsson (2014) distinguishes three levels of liquidity management centralization: fully centralized, centralized with exceptions and highly decentralized. Similarly, Polak shares Potty's four models of establishing the treasury function in international environment, distinguished by degree of decentralization and range of operations undertaken (Potty et al., 2004, as cited in Polak, 2009a, p. 4): full service global, full service local, limited service global and limited service local.

Westerman & Eije (2005) use terms centralization and disintermediation together. Among the benefits they identify cost savings (financing and administration costs, reductions in interest rates etc.), improved monitoring, control and security of cash flows, enhanced efficiency of processes, and finally, in line with Brealey (1992), a decrease of the number of bank accounts open. The disadvantages include incremental costs for information, administration and regulation, distorted attention to local operational needs and disturbed relationships with local banks. Adalsteinsson (2014) identifies the following related indirect costs: interest rate risk (which can be however centralized), cost of maintaining a liquidity buffer and regulatory cost. Polak (2009b) stresses among the benefits of treasury centralization enhanced efficiency, access to information, improved intercompany cash flows, and reduction of various types of costs. On top of that, he adds a dimension of standardisation of processes, which leads into "significant improvements (...) in terms of control and security of cash" (p. 111). Among the factors influencing the centralization decision, Polak (2009b, 2010) includes also company structure and geographical spread with language barriers and different time zones. He shares an idea of building multiple treasury centers to "allow "round the clock" access to the financial markets" (Polak, 2010, p. 89). He perceives centralization of treasury

management as evolution, represented by three phases from phase zero (decentralization), through phase 1 (centralized interest rate risk and FX risk management as well as large liquidity positions; adding cash and liquidity management in phase 2, to full centralization including operating receipts and payments (phase 3) – which may be followed by outsourcing (Polak, 2010, 2011). On a similar topic, Jankensgård (2015) confirmed in his study a hypothesis that usage of FX derivatives creates more value in companies where FX exposure is managed centrally. Another dimension to the topic is brought by Soprano (2015) who distinguishes treasury management function from treasury control function. He states that the operational liquidity risk control is often decentralized, unlike policies and risk setting and monitoring. This opinion is in line with Polak (2011), who distinguishes back office, middle office and front office functions. Soprano (2015) sees an advantage of centralization in optimisation of intra-group liquidity and cost savings not only by means of "having one entry point and interface to the financial markets" (p. 152). Doyle (2006) supports the view of finding the right balance between centralization and decentralization. In his opinion, the knowledge of local conditions and intragroup information sharing forms a competitive asset only when present in combination. From the other perspective, as the biggest advantage of *decentralization* Alonso (2008) identifies transferring the decision rights to managers who are knowledgeable of local information. McCauley et al. (2012) have noticed that in the light of financial crisis, "banks with more centralised or international operations came under intense funding pressure" (p. 7; examples removed by the author), compared to banks with a decentralized multinational structure.

2.2. Economic structure

Drury (2011) comments that "organizations decentralize by creating responsibility centers" (p. 360). Responsibility centers are fundamental elements of responsibility management and according to Král (2018) they serve mainly for application of performance evaluation criteria. Although the categorization of responsibility centers is generally perceived the same way, many modifications are in place (Drobyazko & Yunatskyi, 2019; Fibírová et al., 2019; Schoute, 2008). The traditional classification of department types consists of cost (expense) center, revenue center, profit center and investment center (Drury, 2011; Fibírová et al., 2019; Král, 2018; Schoute, 2008). Drury specifies that "both cost and revenue center managers have limited decision-making authority" (p. 361) as they are responsible for decisions related to controllable inputs (*cost centers*) or financial outputs (*revenue centers*) only. *Profit centers* are centers with increased autonomy, as their managers are responsible for both of the above elements. Similarly to Drury (2011), Ciurlau (2012) presents an additional classification of partial (cost centers and revenue centers) and global responsibility (profit centers). Even higher autonomy have managers of *investment centers*. In Král's (2018) vision, this next step in autonomy is divided between *rentability center* being granted decision rights regarding net working capital, and investment center, which can, in addition, decide about investment projects. Drury (2011) and Schoute (2008) merge these functions into the scope of an investment center. On top of these responsibility centers, a *discretionary expense center* is outstanding. It can be categorized as a separate responsibility center (Král, 2018) or as a subcategory of cost center (Drury, 2011; Fibírová et al., 2019). This center generates long-term benefits while the connected expenses are paid now. As typical examples serve marketing or employee education.

Polák and Kocurek (2007) have touched a similar topic while using terms of cost center, profit center and their combination during a discussion on cash pool. Pokutta & Schmaltz (2011), who conduct a similar research in the banking field, use a term of liquidity center for "a specialized department that produces liquidity at the local interbank market" (p. 627). Inderst & Muller (2003) state that "ceteris paribus, centralization is optimal for projects with a low expected return, or productivity, while decentralization is optimal for projects with a high expected return" (pp. 1034-1035).

3. Results and discussion

The above reviewed approaches to organizational structure are applicable not only to responsibility management, but also to group cash management. To apply the responsibility centers structure into the liquidity management topic, it is necessary to transform the terms and define them in a new way. As resulted from the literature review, in liquidity management field a certain level of centralization is necessary, irrelevant if the central treasury unit is a part of company headquarters, a separate entity, or outsourced. Although discussion about different tools of liquidity management is not in the scope,

Naturally, there is a wide variety of nuances between the extreme scenarios of the case study, and so is the range between "full centralization" and "full decentralization". Referring to the two companies entering the case study, it is clear that liquidity of these two types of group structures shall be managed using different principles. Companies in Group A are independent from any structural funding as their operations resemble a network of separate, individual businesses that are able to fund their own operations. Although their overall independence is subordinate to the group strategy and control, it is anticipated that they enjoy more freedom in decisions than companies in Group B. Group B is structured as a chain of functional companies. As the activities performed by these companies are interdependent, the subsidiaries of Group B are subject to strong need for coordination not only from the cash management perspective, but also by means of organization of the business' operations. Thus, in case of a sales-oriented company in Group B, who is generating cash inflows that need to be further distributed into the group, we will use the term "inflow center". Using analogy, at the subsidiary ("outflow center") in Group B operating as a cost center we can notice regular cash outflows, but only marginal or no inflows. Subsidiaries in Group A can be marked as "self-financing centers" as they generate inflows that are under normal conditions able to cover their cash outflows. The role of intercompany cash flows is in Group A similar to flows with external business partners.

Two elements of liquidity management will be further distinguished: operational liquidity management and investments of excess cash. Obviously, Group B shows strong need for centralized liquidity management to fund operations of the outflow centers by withdrawing excess cash from the inflow centers. This requirement is noticeably weaker for the self-financing centers in Group A. More difficult is the centralization question of the investment function. For this paper, meaning of an investment center needs to be transformed and matched to the level of centralization as well. Only the treasury center, where the cash is collected from the group, can make investments in major scope to fully use the benefit of economies of scale. Under some circumstances, this setup is also possible for inflow centers. To distinguish this high-scope investing from subsidiaries being granted the right to freely invest their excess funds, it is necessary to use terms "low-scope investment center" and "high-scope investment center". The latter one can be omitted while referring to the treasury center.

Table 1 – Matrix of cash flow centers for investment purposes

Type of cash center	Not an investment center	Low-scope investment center	High-scope investment center
Outflow center	Possible	Possible (conditional)	Never
Inflow center	Possible	Possible	Possible (conditional)
Self-financing center	Possible	Possible	Never
Treasury center	Never	Never	Always

The above matrix represents the possibilities of the company structure and level of liquidity management centralization by means of *decentralized* financial investment function. Combination of an outflow center investing in low-scope and an inflow center investing in high-scope are marked as possible, but conditional. Dependent on the company's decision on frequency and volume of funds

and injections transferred to the outflow center, the center might be granted a right to invest the very short-term excess of cash to rather liquid financial instruments. As mentioned above, the inflow center may be merged with a treasury center, if that is relevant for the specific company.

However, a problem arises in Group A, whose companies are rather independent in the cash question. As each of the center is holding a precautionary cash balance, the total group liquidity is distorted within the subsidiaries and the group loses the benefit of economies of scale for investment purposes. A cost vs. benefit analysis must be performed by the company to reflect its specific conditions when deciding on the need of managing this liquidity centrally by using intercompany liquidity tools (such as cash pool). Variables into this analysis are interest rates in different countries, foreign exchange rates, bank fees for transfers, fees of running liquidity tool, balance monitoring and administration fees as well as opportunity costs for the time while cash is in transit, especially within countries that apply currency control. Compliance with tax and legal regulation needs to be considered as well. The need for centralization of the cash flow centers is summarized in the below table. Treasury center is not part of this table as it serves as the highest level of centralization.

Table 2 – Need for liquidity centralization by element

Type of cash center	Operational liquidity	Financial investments
Outflow center	Strong	Strong
Inflow center	Strong	Rather weak
Self-financing center	Weak	Rather strong

After deciding on the degree of centralization, further decisions must be made by the management, followed by preparation of proper guidelines. These decisions give answers to the following questions: In which intervals to withdraw from inflow centers and to fund outflow centers? Shall investments to financial instruments take place in the inflow center directly or shall they be transferred at the earliest convenience into the entity, where the liquidity management is centralized?

3.1. Conclusion

The overall aim of this paper is to offer another view into the discussion on appropriate degree of liquidity management centralization by adding an interdisciplinary dimension. To reach this aim, an analysis of literature sources was performed, followed by a case study to answer the research question: How can economic structure, represented by responsibility centers, be translated into centralization of corporate liquidity management?

Summarizing the literature review on the cash management field, the authors have come to an agreement that at least partial centralization of liquidity management is the preferred option. A continuous trend of centralizing liquidity management has been identified by multiple authors. Benefits of centralization can be summarized into the categories of cost reduction and improvements in efficiency, control and security. The disadvantages include incremental costs of coordination, regulatory and legal compliance and a risk of distorted attention to local needs.

An interdisciplinary dimension was added by transforming responsibility centers into liquidity centers. To merge and apply these principles, case studies of Groups A and B were used. These groups represent two different organizational structures and demonstrate their possible structural dependence on intercompany cash flows. Two elements of liquidity management were distinguished: operational liquidity management and investments of excess cash. Combinations of inflow center, outflow center and self-financing center on operational side, and low-scope investment center, high-scope investment center and "not an investment center" on the investment side are presented in a matrix which demonstrates the possibilities of the company structure and level of liquidity management centralization by means of decentralized financial investment function. The issue of

holding precautionary cash balances resulting in losing the benefit of economies of scale of investments is addressed in another table, which summarizes the need for liquidity centralization by element. However, a cost vs. benefit analysis must be performed by the company to reflect its specific conditions when deciding on the degree of its treasury management centralization.

Due to the limited scope of this paper, the classification of internal and external drivers for liquidity centralization is not discussed further. Along with an analysis of the resulting threats and opportunities it would be a vital topic for the subsequent research.

4. Acknowledgement

5. References

ACCA, & Kaplan Publishing UK. (2019). Advanced Performance Management (APM). Kaplan Publishing UK.

Adalsteinsson, G. (2014). The liquidity risk management guide: From policy to pitfalls. Wiley.

Alonso, R., Dessein, W., & Matouschek, N. (2008). When Does Coordination Require Centralization? American Economic Review, 98(1), 145–179. https://doi.org/10.1257/aer.98.1.145

Barbas, J. C. (2009). The Single Euro Payments Area: A strategic business opportunity. Journal of Corporate Treasury Management, 2(3), 246–251.

Blažek, L., Masarykova univerzita, & Centrum výzkumu konkurenční schopnosti české ekonomiky. (2010). Nadnárodní společnosti v České republice I.: (Empirické studie). Masarykova univerzita.

Brealey, R. A., Myers, S. C., Melnyčok, J., Tůma, M., & Tůma, Z. (1992). Teorie a praxe firemních financí. Victoria Publishing.

Christie, A. A., Joye, M. P., & Watts, R. L. (2003). Decentralization of the firm: Theory and evidence. Journal of Corporate Finance, 9(1), 3–36. https://doi.org/10.1016/S0929-1199(01)00036-0

Ciurlau, L. (2012). Measuring Financial Performance Types of Responsibility Centers. Anale. Seria Stiinte Economice. Timisoara; Timisoara, 18(suppl.), 332–336.

Cobben, H. (2008). How payment factories improve liquidity management through centralised control and visibility. Journal of Payments Strategy & Systems, 2(2), 192–196.

Doyle, D. P., Wagner, J., Matyáš, O., & Menšík, M. (2006). Strategické řízení nákladů = Cost Control—A strategic guide. ASPI.

Drobyazko, S., & Yunatskyi, M. (2019). Evaluation of Effectiveness of Responsibility Centers in the Management Accounting System. Academy of Accounting and Financial Studies Journal, Volume 23.

Drury, C. (2011). Cost and management accounting (7. ed). Cengage Learning.

Fibírová, J., Šoljaková, L., Wagner, J., & Petera, P. (2019). Manažerské účetnictví: Nástroje a metody (3rd ed.). Wolters Kluwer.

Gitman, L. J., Moses, E. A., & White, I. T. (1979). An Assessment of Corporate Cash Management Practices. Financial Management, 8(Spring 1979), 32–41.

Inderst, R., & Muller, H. M. (2003). Internal versus external financing: An optimal contracting approach. Journal of Finance, 58(3), 1033–1062. https://doi.org/10.1111/1540-6261.00557

Jankensgård, H. (2015). Does Centralisation of FX Derivative Usage Impact Firm Value?: Does Centralisation of FX Derivative Usage Impact Firm Value? European Financial Management, 21(2), 309–332. https://doi.org/10.1111/j.1468-036X.2013.12014.x

Jensen, M. C., & Meckling, W. H. (1995). Specific and General Knowledge, and Organizational Structure. Journal of Applied Corporate Finance, 8(2), 4–18. https://doi.org/10.1111/j.1745-6622.1995.tb00283.x

Král, B. et al. (2018). Manažerské účetnictví (4. rozšířené a aktualizované vydání). Management Press.

McCauley, R., McGuire, P., & von Peter, G. (2012). After the global financial crisis: From international to multinational banking? Journal of Economics and Business, 64(1), 7–23. Scopus. https://doi.org/10.1016/j.jeconbus.2011.06.004

Milkau, U., & Bott, J. (2015). Digitalisation in payments: From interoperability to centralised models? Journal of Payments Strategy & Systems, 9(3), 321–340.

Pokutta, S., & Schmaltz, C. (2011). Managing liquidity: Optimal degree of centralization. Journal of Banking & Finance, 35(3), 627–638. https://doi.org/10.1016/j.jbankfin.2010.07.001

Polak, P. (2009a). Location Criteria for Establishing the Treasury Centres in South East Asia (SSRN Scholarly Paper ID 1471889). Social Science Research Network. https://doi.org/10.2139/ssrn.1471889

Polak, P. (2009b). 'The Centre Holds'—From the Decentralised Treasury Towards Fully Centralised Cash and Treasury Management. Journal of Corporate Treasury Management, 3(2), 109–112.

Polak, P. (2010). Centralization of Treasury Management in a Globalized World (SSRN Scholarly Paper ID 1702687). Social Science Research Network.

Polak, P. (2011). Global challenges in corporate finance and treasury management—Centralisation and internationalisation. Journal of Corporate Treasury Management, 4(3), 242–250.

Polak, P., & Kocurek, K. (2007). Cash and Working Capital Management in The Czech Republic. Investment Management & Financial Innovations, 4, 17–30.

Reed, W. L. (1972). Profits From Better Cash Management. Financial Executive, 40(May 1972), 40–56.

Regino, L., & Bernardelli, S. (2008). Using technology to build a nimble and proactive treasury organisation. Journal of Corporate Treasury Management, 2(2), 175–179.

Roun, V. (2007). Transfer prices within business combinations. Český Finanční a Účetní Časopis, 2007(2), 17–27. https://doi.org/10.18267/j.cfuc.218

Schoute, M. (2008). Determinants of Responsibility Centre Choices – An Empirical Study at the Managerial Level. 6(1), 12.

Smith, K. V., & Sell, S. B. (Eds.). (1980). Readings on the management of working capital (2d ed). West Pub. Co.

Soenen, L. A. (1986). International cash management: A study of the practices of U.K.-based companies. Journal of Business Research, 14(4), 345–354. Scopus. https://doi.org/10.1016/0148-2963(86)90025-1

Šoljaková, L. (2009). Strategicky zaměřené manažerské účetnictví. Management Press.

Soprano, A. (2015). Liquidity management: A funding risk handbook. Wiley.

van Alphen, W., & Huiskes, A. (2008). Global liquidity management: From vision to reality. Journal of Corporate Treasury Management, 1(4), 326–332.

Venkat, S., & Baird, S. (2016). Liquidity risk management: A practitioner's perspective. Wiley.

Wagner, J. (2009). Měření výkonnosti: Jak měřit, vyhodnocovat a využívat informace o podnikové výkonnosti. Grada.

Wessels, B., & Van Rooyen, J. H. (2010). Bank strategic and risk management: A survey to determine the impact of the current world financial crisis on future bank management in South Africa. Corporate Ownership and Control, 8(1 D), 390–401. Scopus. https://doi.org/10.22495/cocv8i1c3p6

Westerman, W., & von Eije, H. (2002). Multinational cash management and conglomerate discounts in the euro zone. International Business Review, 11(4), 453–464. https://doi.org/10.1016/S0969-5931(02)00019-7

Westerman, W., & von Eije, H. (2005). Multinational Cash Management in Europe Towards Centralisation and Disintermediation: The Philips Case. Managerial Finance; Patrington, 31(10), 65–74. http://dx.doi.org.zdroje.vse.cz:2048/10.1108/03074350510769929

IMPACT OF BUSINESS INTELLINGENCE TOOLS ON THE QUALITY OF MANAGEMENT ACCOUNTING

Libuše Šoljaková

University of Economics, Prague
soljak@vse.cz

Keywords

Business Intelligence, Management Accounting, Reporting

Abstract

Many organizations implement BI technologies to support management accounting. This paper focuses on BI usage in Czech organizations both as a part of management accounting tasks and techniques and also in improving the quality of management accounting information. Methodologically, the data were collected via a questionnaire, then statistically analysed. The results of this study support the hypothesis that BI is widely used in the Czech Republic and has improved the quality of management accounting information.

1. Introduction

Many organizations are implementing business intelligence (BI) technologies to support management accounting (particularly the decision-making and controlling processes). Management accounting has clear links to and should benefit from the application of BI technologies. This paper studies the use of BI in supporting decision-making, reporting and controlling processes and the impact of BI on the quality of such processes.

Most research has focussed on how the adoption of ERP systems (or integrated information systems – information base and support for BI technologies) affects management accounting tasks and techniques. Although these systems have no doubt increased the efficiency of collecting and reporting accounting data (Cooper and Kaplan, 1997; Davenport, 1998), many have concluded that they often have stabilising effects on management accounting practices rather than engendering any direct or significant changes (Granlund and Malmi, 2002; Rom and Rohde, 2007; Scapens and Jazayeri, 2003). Nevertheless, ERP systems enable better control (Chapman and Kihn, 2009; Granlund and Malmi, 2002; Rom and Rohde, 2007; Scapens and Jazayeri, 2003; Wagner et al., 2011) and positively affect organizational performance (e.g., Hunton et al., 2003; Nicolaou, 2004; Nicolaou and Bhattacharya, 2006; Velcu, 2007).

Also, studies have looked at the changing role of the management accountant because of the digitization of accounting processes. While management accountants' roles have become more business-oriented and strategic, almost consultant-like (Granlund and Malmi, 2002; Quattrone and Hopper (2001)), better access to information has led to the decentralization of the management accounting function, whereby management accounting tasks are increasingly performed within other elements of an organization (Caglio, 2003; Rom and Rohde, 2007).

2. Literature Review

BI is defined as a technology and process for analysing data and presenting actionable information to help organizational decision makers make better decisions (Chaudhuri et al., 2011; Howson, 2007; Davenport, 2006, 2010, 2013, 2014; Sharda et al., 2014). BI is an "umbrella term" in that it encompasses a variety of technologies and methodologies that enable organizations to collect data from internal and external sources, prepare it for analysis, develop and run queries against the data, and create reports, dashboards and data visualizations to make the results available to end users (Rikhardson and Yigitbasioglu, 2018).

Previous research into BI supports several management accounting tasks, such as cost forecasting, product profitability analysis, the financial impact of production changes, and assessments of customer segment profitability (Bronzo et al., 2013; Lee and Park, 2005). The use of BI-supported analytical methods improves organizational performance in several dimensions, including financial, customer, process and learning & growth (Bronzo et al., 2013). Many BI&A software solutions are also available for planning, reporting and consolidation tasks (Marx et al., 2012). Both technical characteristics and application characteristics determine the level of maturity of BI application in planning, reporting and consolidation in companies. The five-stage maturity model developed by Marx et al. (2012) can be used to map a company's BI maturity regarding these tasks and plan further developments. BI can be used in making different types of routine and non-routine decisions, including product pricing decisions and the choice of product mix (Kowalczyk and Buxmann, 2015). However, there are tensions in implementing data-centric decision making in organizations. These include tensions between the flexibility and stability of methods and data sources, tensions involving the complexity and understandability of analysis methods, and tensions between broad and focused scopes of analyses. A number of tactics can be used to achieve ambidexterity and resolve these issues.

On the other hand, BI is criticised for being too passive. Situated systems for decision support are complete solutions in the sense that they are active: they can continuously sense the environment, suggest responses, support the implementation of the decision and provide subsequent monitoring. A study by Vahidov and He (2010) provides evidence for the superiority of such solutions compared to conventional decision-support systems in a personal finance context (Vahidov and He, 2010). Management accountants in this respect can help in promoting the integration of BI into decision making.

Rikhardson and Yigitbasioglu (2018) suggested the framework below (Fig. 1) based on elements of BI and MA. The five themes identified from the literature review combine various elements of both BI and MA.
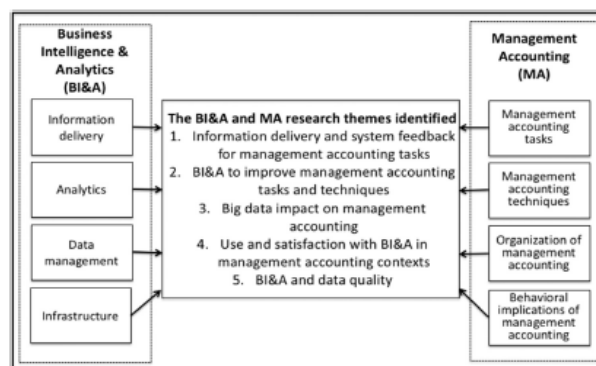
Fig. 1; Source: Rikhardsson,P., Yigitbasioglu, O., 2018. Business intelligence & analytics in management accounting research: Status and future focus. International Journal of Accounting Information Systems.29 (2018) 37-58

This paper focuses only on the second issue, "BI to improve management accounting tasks and techniques", and the fifth issue, "BI and data quality". Based on these conclusions we have developed the following hypotheses:

H1: Organizations in the Czech Republic widely use BI for support in decision-making, reporting and controlling processes.

H2: BI improves the quality of decision-making, reporting and controlling processes.

3. Data gathering, sample

The data were collected via a pilot questionnaire survey developed at the beginning of 2020.

The questionnaire consisted of questions exploring the application of BI in connection with basic management accounting tasks and techniques, plus the impact of BI on the improvement of the quality of information delivered by management accountants. The majority of questions were answered on a seven-point scale (1 definitely disagree and 7 definitely agree). Some respondents (total 8) we also contacted directly and asked four additional questions.

The sample of companies for our research was sourced from the Albertina CZ Gold Edition database. The companies selected were domiciled in the Czech Republic, had more than 50 employees, and boasted a turnover above 256 million CZK. Our research was cross-industrial, but the following industries were excluded (according to NACE ver. 2) 'O – Public administration and defence', 'P – Education', 'Q – Human health and social work activities', 'S – Other service activities', 'T – Activities of households as employers', and 'U – Activities of extraterritorial organisations and bodies'. We randomly selected 200 companies matching these criteria.

The data were gathered from January 2020 to February 2020 and we obtained, in total, 20 responses.

4. Results and discussion

In our research we have investigated the use of BI for management accounting support. We have analysed the application of BI in the following areas:

- Process planning, forecasting and variance analyses
- Profitability analyses (product and customer segments)
- Decision making (routine and non-routine)
- Reporting - traditional (tables and graphs) and modern (dashboard and interactive presentation)

According to the results of the research, BI is widely applied in planning, controlling, reporting and decision making. Profitability analysis (customer segment and product) and traditional reporting achieved the highest usage of BI. On the other hand, modern reports are not widely used, probably because most managers prefer traditional reports based on tables and graphs to modern visualisation techniques.

BI is also important for planning and variance analysis and routine decision making. According to interviews with respondents, it allows the preparation of budgets and forecasts in much more precise and detailed ways whilst at the same time allowing one to take more variables into account. The lowest use is in non-routine decision making. This is mainly due to the fact that the systems currently in use are not yet at a level sufficient to deal with random, unique and non-repetitive user requirements.

Descriptive statistics related to the use of BI can be found in Table 1.

Table 1 – Descriptive statistics of BI use

Description	Average	Median	Mode	Std. deviation
Forecasting and planning	4.05	4	7	2.109
Variance analysis	3.85	4	4	1.824
Product profitability analysis	5.00	5	5	1.378
Customer segment profitability	5.45	5.5	7	1.396
Routine decision making	4.90	5	7	1.700
Non-routine decision making	2.35	2	1	1.195
Reporting - traditional	5.10	5	4	1.338
Reporting – modern	2.60	3	3	1.241

Our computations also confirmed hypothesis H1 concerning the expected increase of BI usage.

The quality of accounting information was measured according to:

- Correctness of data processed subsequently into output information used in managerial control

- Information relevance (focus on important information for decision making or control)

- Timeliness of information delivery

- Formal level of rendered information

- Neutrality and objectivity of information

- Information flexibility – open to the adoption of changes and variables

- Information comparability and consistency – during the time period and across each organisation

Regarding information improvement, it was shown that management accountants confirm a better quality of information when using BI. Our findings upon the correctness of calculations and information flexibility are absolutely clear. There is also significant improvement in indicators of information comparability and consistency, the formal level of information, and the timeliness of rendered information. All respondents confirm that BI eliminates formal and calculation errors and value the easy recalculation arising from changes in input variables (exchange rates, interest rates, product selling price, purchasing costs etc.). Only limited benefits were found in terms of neutrality and objectivity, respondents feeling that even without BI they were able to provide objective and neutral information.

Nevertheless, the success of the system is contingent on both the system itself and on the user's knowledge of the system. Thus, user training in BI&A functionality in management accounting is essential (Deng and Chi, 2012). Exploring this in the context of management accounting constitutes a key research topic, given the importance of management accounting in decision making.

Descriptive statistics related to the hypothesis, "BI improves quality of management accounting information", are found in Table 2.

Table 2 – Descriptive statistics of quality improvement.

Description	Average	Median	Mode	Std. deviation
Correctness of calculation	6.45	6	6	0.497
Information relevance (focus on important information)	3.65	3	2	1.682
Timeliness of rendered information	4.30	4	4	1.005
Formal level of rendered information	4.70	5	4	1.487
Neutrality and objectivity of information	2.75	3	4	1.090
Information flexibility	6.45	6	6	0.497
Information comparability and consistency	4.90	5	3	1.578

5. Conclusion

Our paper uses data from an original empirical survey conducted in the Czech Republic at the beginning of 2020. Primarily, the use of BI for management accounting support and its impact on the quality of management accounting information were investigated.

We found that BI is widely applied in the planning, forecasting and controlling process, as well as for routine decision making and reporting, with less use BI in non-routine decision making. .

We found that BI possibly improves the quality of management accounting information in terms of its flexibility, correctness of calculation, comparability and consistency, and timeliness.

6. Acknowledgement

7. References

Bronzo, M., de Resende, P.T.V., de Oliveira, M.P.V., McCormack, K.P., de Sousa, P.R., Ferreira, R.L. (2013). Improving performance aligning business analytics with process orientation. International Journal of Accounting Information Systems 33 (2), 300–307.

Caglio, A. (2003). Enterprise resource planning systems and accountants: towards hybridization? European Accounting Review. 12, 123–153.

Chapman, C.S., Kihn, L.-A. (2009). Information system integration, enabling control and performance. Accounting Organizations and Society. 34 (2), 151–169.

Chaudhuri, S., Dayal, U., Narasayy, V. (2011). An overview of business intelligence technology. Commun. ACM 54, 88–98.

Cooper, R., Kaplan, R.S. (1997). The promise–and peril–of integrated cost systems. Harvard Business Review. 76 (4), 109–119.

Davenport, T. (1998). Putting the enterprise into the enterprise system. Harvard Business Review 76, 121–131.

Davenport, T. (2006). Competing on analytics. Harvard Business Review 84, 98–107.

Davenport, T. (2010). BI&A and organizational decisions. International Journal Business Intelligence Research. 1, 1–12.

Davenport, T. (2013). Analytics 3.0. Harvard Business Review 91, 64–72.

Davenport, T. (2014). Big Data at Work: Dispelling the Myths, Uncovering the Opportunities. Harvard Business Review Press, Harvard.

Deng, X., Chi, L. (2012). Understanding postadoptive behaviours in information systems use: a longitudinal analysis of system use problems in the business intelligence context. Journal of Management Information Systems. 29 (3), 291–326

Granlund, M., Malmi, T. (2002). Moderate impact of ERPS on management accounting: a lag or permanent outcome? Management Accounting Research 13, 299–321.

Howson, C. (2007). Successful Business Intelligence: Secrets to Making BI&A a Killer App. McGraw-Hill, New York.

Hunton, J.E., Lippincott, B., Reck, J.L. (2003). Enterprise resource planning systems: comparing firm performance of adopters and nonadopters. International Journal of Accounting Information Systems. 4 (3), 165–184.

Kowalczyk, M., Buxmann, P. (2015). An ambidextrous perspective on business intelligence and analytics support in decision processes: insights from a multiple case study. Decision Support Systems 80, 1–13.

Lee, J.H., Park, S.C. (2005). Intelligent profitable customers segmentation system based on business intelligence tools. Expert Systems with Applications. 29 (1), 145–152.

Marx, F., Wortmann, F., Mayer, J. (2012). A maturity model for management control systems - five evolutionary steps to guide development. Business & Information Systems Engineering4 (4),193–207.

Nicolaou, A.I. (2004). Firm performance effects in relation to the implementation and use of Enterprise resource planning systems. Journal of Informaton Systems. 18 (2), 79–105.

Nicolaou, A.I., Bhattacharya, S. (2006). Organizational performance effects of ERP systems usage: the impact of post-implementation changes. International Journal of Accounting Information Systems 7 (1), 18–35.

Quattrone, P., Hopper, T. (2001). What does organizational change mean? Speculations on a taken for granted category. Management Accounting Research .12 (4), 403–435.

Rikhardsson,P., Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. International Journal of Accounting Information Systems 29 37-58

Rom, A., Rohde, C. (2007). Management accounting and integrated information systems: a literature review. International Journal of Accounting Information Systems 8, 40–68.

Scapens, R.W., Jazayeri, M. (2003). ERP systems and management accounting change: opportunities or impacts? A research note. European Accounting Review. 12 (1), 201–233.

Sharda, R., Delen, D., Turban, E., (2014). Business Intelligence and Analytics: Systems for Decision Support, 10th ed. Prentice Hall, New York.

Vahidov, R., He, X.(2010(. Situated DSS for personal finance management: design and evaluation. Information and. Management. 47 (2), 78–86.

Velcu, O. (2007). Exploring the effects of ERP systems on organizational performance: evidence from Finnish companies. Industrial Management & Data Systems.. 107 (9), 1316–1334.

Wagner, E.L., Moll, J., Newell, S. (2011). Accounting logics, reconfiguration of ERP systems and the emergence of new accounting practices: a sociomaterial perspective. Management Accounting Research. 22 (3), 181–197.

DIGITAL SECURITY, SAFETY, TRUST OF SMART SYSTEMS – FOR A RESILIENT

TOWARDS A RESILIENT SOCIETY – TECHNOLOGY 5.0, RISKS AND ETHICS

Erwin Schoitsch

AIT Austrian Institute of Technology GmbH (Vienna)
erwin.schoitsch@ait.ac.at

Keywords

Smart Systems, Resilient Society, Sustainable Society, Business Resilience, Autonomous Systems, Industry4.0, Systems of Systems, Cyber-physical systems, Internet of Things (IoT), Artificial Intelligence, Society 5.0, Ethical Aligned Design, Sharing Society, Sharing Economy.

Abstract:

Society 5.0 as "super-smart society" is the key element of the Japanese 5th Science and Technology Basic Plan by the Council for Science, Technology and Innovation 2016. It became a political highlight of the Japanese government and was taken over 2017 and 2018 as a vision for the Japanese economy and society, to take over the lead ahead of the world to make people's life more comfortable and sustainable.

Smart Systems, i.e. largely deployed and interconnected CPS (cyber-physical system and IoT networks) and integrated intelligence and autonomy are considered the drivers of innovation. In all industrial and social areas highly automated or autonomous intelligent systems are taking over tasks and services – and maybe, one day, control of our lives. The keynote will raise questions and discuss impact, risks, ethical issues and challenges such as "Can a technology dependent and technology driven society be resilient and sustainable? Can technology make a society resilient and sustainable? Will the role of humans change in such a society? What are the trade-offs with respect to human rights, self-determination, independence or will "Big Brother" control risks become overwhelming? The keynote will address issues that are already evident now and how resilience, sustainability and ethical issues are now discussed in different context – particularly how can a resilient society manage a crisis like the Climate Crises, and Covid-19 – a situation that has revealed vulnerabilities and will hopefully lead to a rethinking of some economic and societal systemic issues.

1. Introduction and Overview – Smart Systems Everywhere

Our society is facing many challenges now – upcoming crisis in many respects, e.g. the climate crisis, the Covid-19 (Corona) crises, and economic as well as social crisis impacts in the future. In many parts of the world, we have exaggerated usage of our natural resources in a ruthless manner often just for short term profits and benefits of a few. This situation is encompassed by an enormous progress in technology, particularly in electronic systems and materials. Smart systems are already everywhere –but integrated into "super-smart systems-of-systems" with advanced algorithms and artificial intelligence and "big data" in the background. The pretended purpose of this high degree of connectivity is to serve customers (peoples) needs for their benefit – but the impact is much higher. Challenges are not only on the technical and usability level, they are on the level of increased risks, threats and vulnerabilities of the systems concerning safety, security and privacy. Last but not least

are we afraid of the "transparent citizen", the total surveillance ("Big Brother") like shown in a much simpler and less professional manner in the famous book "1984" by George Orwell; less professional because of lack of the technology that is easily available now. Even the attacks on human rights, freedom and independence are now much "smarter", and most of us are often not aware of the degree of surveillance and control, even in our democracies in the so-called "Western World".

This will impact considerably our lives and lifestyle, and, as a consequence, we will have to face the new challenges, opportunities and risks. Not to forget the economic risks – recent crisis like Covid-19 lock-downs have demonstrated our vulnerability because of the dependency on worldwide functioning supply chains.

The technological basis is laid by IoT (human – smart devices interaction and communication) and IIoT (IoT in industrial context, machine-to-machine communication) as infrastructure (connectivity), and CPS (Cyber-physical Systems) as "things" or "devices". But smartness is more – it means intelligence, cognitive systems and technology, machine learning and artificial intelligence, security, big data and cloud connectivity, involving many domains of everyday life and digital transformation of our world.

With respect to safety, cybersecurity, privacy, and data sovereignty, the new "cognitive technologies" are a severe concern for specialists, politicians and citizens, and raise severe ethical and societal concerns, and dependency on these technologies in a networked world is an additional concern now.

For our future it is most important now to strive for a "resilient society" and a "sustainable economy", both relying on each other. Such a strategy will cope with the challenges particularly by the of the crises mentioned before – climate, pandemic and economic risks, stability of society and long-term survival and peace.

2. The Role of Technology: Enabler for a Resilient Society or Risk?

In all technological and industrial advanced regions of our world, research and innovation in smart, intelligent systems and systems-of-systems are considered essential and a lot of money is invested by governments and industry. On European level, organizations like AIOTI [AIOTI], the Alliance for Internet of Things Innovation, which takes care of the IoT aspects in 13 Working Groups, or the industrial associations ARTEMIS [ARTEMIS] (Advanced Research and Technology on Embedded Intelligent Systems), EPoSS [EPoSS] (European Technology Platform for Smart systems Integration) and AENEAS (Association for European Nano-Electronics Activities), which are the private partners in the ECSEL Joint Undertaking, a European PPP within Horizon 2020 (Public-Private Partnership) with an industry-oriented Research Program, and other PPPs, take care of further development of research, standardization and promotion of these topics, together with the European Commission and national funding authorities. China is already keeping up with Europe, US and Japan, e.g. with its AI initiative and strategy.

The digital transformation of European business and society is a major goal of the EC. EC Growth, the DG (Directorate General) for Internal Market, Industry, Entrepreneurship and SMEs, considers digital transformation as a key element for European growth, because Europe can build on its strength in traditional sectors and can take up the potential and challenges of advanced digital technologies. Technologies considered in this context are IoT, big data, advanced manufacturing, robotics, 3D printing, blockchain technologies and artificial intelligence (see European Commission, 2018 - 2020).

The initiative "Digitizing European Industry" targets to meet Europe's needs to join forces under a common strategy that takes digitalization of the EU's economy forward in order to unlock the full potential of the 4th industrial revolution. The pillars of this initiative are:

Figure 1: Pillars of the European Initiative "Digitizing European Industry"

(Source: https://ec.europa.eu/digital-single-market/en/pillars-digitising-european-industry-initiative)

In the booklet "My agenda for Europe" of Ursula von der Leyen, the new President of the European Commission, is one chapter dedicated to "A Europe fit for the digital age". It focuses on AI, IoT, 5G, and ethical and human implications of these technologies, empowering people through education and skills, and on protecting ourselves with respect to the risks of these technologies.

But it is not only economics and competitiveness of European industry – the recently announced "Green Deal" as basis and justification for the upcoming huge efforts are targeting at "Resilience" and "Sustainability" in all technology-driven areas, like manufacturing, mobility, public and industrial administration, particularly considering also societal and ethical aspects.

Additionally, DG Growth delivers an annual report on standardization, e.g. the "Rolling Plan on ICT Standardization", which includes most of the relevant areas in this paper's context and is a key pillar in Digitalization, and have started a Joint Initiative on Standardization (JIS) http://ec.europa.eu/growth/single-market/europeanstandards/notification-system_en, although they do primarily consider the European SDOs (Standardization Organizations, ESOs) CEN, CENELEC and ETSI. The same is done by the international standardization organizations ISO and IEC, particularly in their Joint Technical Committee JTC1 (Information Technology), which cover all these topics of "Joint interest", particularly IoT (SC41), AI (SC42) and Security (SC27) and Software and Systems engineering (SC7).

The straightforward goal of these efforts is to have advanced technologies in ICT end ECS (Electronic components and systems) the enable a sustainable and resilient society in the economic and human factors and well-being sense. As examples like the movement of "Resilient Cities" and "Sustainable Growth" show, these technologies are a chance to overcome current waste of resources and to optimize resource usage, recycling and reuse sustainable production not only in industry but also in agriculture (smart farming, in large cities even "vertical farming" in an until now unseen performance making large population agglomerates less vulnerable to supply risks in case of a crisis), mobility, work spaces and living spaces ("smart buildings", "smart homes"). "Business resilience" is also a key objective, e.g. the COVID-19 impact on global service delivery models, as several studies have described (SSON-Study "Reality Bites – Impact of COVID-19 on Global Service Delivery Models"").

On the other hand, human freedom, independence and privacy, as well as resilience are certainly endangered, the latter because of too much dependency on technologies when you have not nearby access concerning maintenance and update.

Examples how AI can influence our job situation is demonstrated by measures of some large companies, and even public organizations, particularly by biased training datasets (taken from newsletters):

- Microsoft replaces journalists by AI systems (for MSN-website News)
- Public unemployment service uses AI algorithms to select clients: elder women skipped out of the search for jobs (without individual check of qualification)
- Racist decisions by AI systems (court, police, gender issues)

Therefore, we have to look not only at system properties like safety, security or privacy, but also on other aspects, how we can "TRUST" these technologies.

3. Crisis as a challenge and chance

In a publication "Why European companies need to complete the digital transformation to remain viable and competitive", this crisis is considered as a "turbo" to advance faster in our digitalization – to become more resilient and sustainable as economy and society.

The Corona-crisis is looked at as a "disruptive force and digitalization turbo" (citation):

"Breakthrough for digitalization", "How the virus makes digitalization move forward" or "Corona makes the Internet a matter of course" – this is how the headlines of articles describing the effects of the coronavirus on the digitalization of various areas of life are currently read.

- **Home office:** Almost all companies allow flexible working from home; video conferencing is becoming a matter of course.
- **Digital classroom:** As schools, kindergartens and universities are closed for the time being, eLearning is becoming a mass phenomenon and is currently being experimented a lot – even if the "digitalization gap" in areas of schools is still blatant.
- **Online medicine:** Information and diagnosis are shifting almost massively to digital channels. New forms of care and support are emerging.
- **Entertainment & culture:** Concerts and events are taking place virtually; online streaming is causing the biggest peak in internet traffic ever measured at the Frankfurt exchange node DE-CIX: 9.1 Terabit/s!
- **Social life:** In times of "social distancing", digital media and the Internet play an immanently important role for private and social networking and social solidarity. Thus even the problems with "fake news" fade into the background.
- **Digital town hall:** Authorities and public institutions are restricting public access and expanding digital contact points and services.

The conclusion ends in a recommendation for a future-proof corporate and IT strategy 2030: "Digital resilience and hybrid organisation", which is figured out in Figure 2 (from "Crisp © research by Cloudflight"):
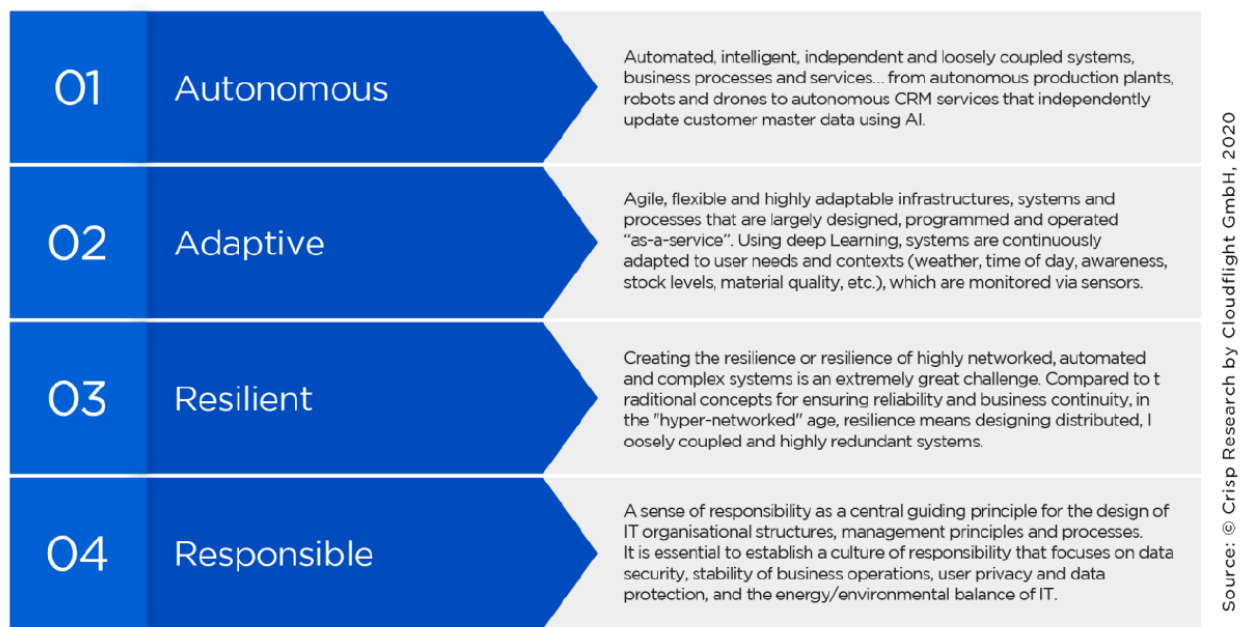
Figure 2: Digital resilience for companies and organizations

4. Trustworthiness of Smart Systems and Systems-of-Systems

With respect to "Resilience" and "Sustainability", SC7, SC41 and SC42 join forces as enablers by creating the adapted system property "Trustworthiness". This is particular of importance if we look at the requirements for a resilient, sustainable society and economy – which goes beyond the classical technical requirements.

"Trustworthiness" as a term was created to include more than the established system properties of dependability (safety, security, reliability, availability, maintainability), but also other stakeholder interests. One of the rather complete definitions we found was:

What is trust?

Trust: degree to which a user or other stakeholder has confidence that a product or system will behave as intended <ISO/IEC 25010:2011(en) Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models (= reliable and dependable)>

ISO/IEC JTC1 AG7 Trustworthiness Definition (July 2019, source ISO/IEC JTC1 AG7 presentation by François Coallier):

- Trustworthiness corresponds to the ability to meet stakeholders' expectations in a verifiable way.

- Depending on the context or sector, and also on the specific product or service, data, and technology used, different characteristics apply and need verification to ensure stakeholders expectations are met.

- Characteristics of trustworthiness include, e.g. reliability, availability, resilience, security, privacy, safety, accountability, transparency, integrity, authenticity, quality, usability.

- Trustworthiness is an attribute that can be applied to services, products, technology, data and information as well as, in the context of governance, to organizations.

- Trustworthiness is ensured and maintained through a sound governance framework and systems engineering practices.

- Trustworthiness can contribute to the building of confidence.

The IIoT Group (Industrial Internet of Things), although traditionally biased by the purely technical and economic aspects, talks already care of these extended requirements (Figure) from their Newsletters) and define as a minimum:
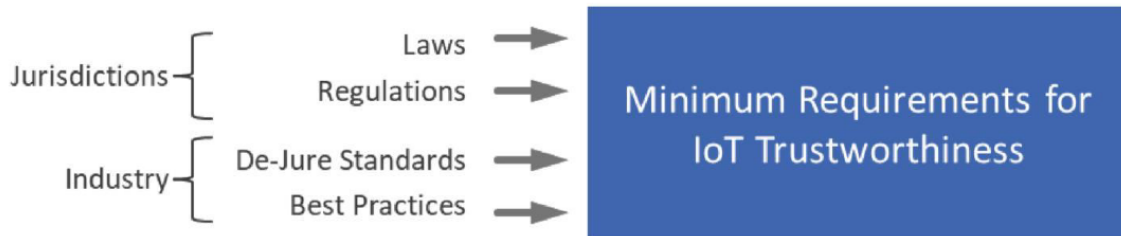
Figure 3: Minimum Requirements as defined by external parties

The societal and political part of resilience and sustainability is defined by the public stakeholders' interests who provide the framework for a resilient and sustainable implementation ("Jurisdiction").

But what about human society? – The Japanese Strategy Society 5.0

5. Society 5.0 – Aiming for a New Human-Centred Society

There are far reaching concepts implemented around us with the "Smart Systems Everywhere". On the other hand, all technologically and economically highly developed countries face chronic social challenges, which add up to the challenges and risks impacted by all the "4.0" revolutions to people and society.

The Japanese strategy "Society 5.0" envisions a so-called "Super-Smart Society". It should create a sustainable, inclusive socio-economic system, powered by the achievements of the fourth industrial revolution and the digital technologies. It leads far beyond Industry 4.0, it aims at benefitting the whole society by utilizing the integration of cyberspace and physical space [UNESCO 2019]. It is fully supported by government and industry to revitalize Japan, an extremely aging society with shrinking work force [Japan Business Federation, 2016, 2017] [Japanese Government, 2017, 2018].

Evolution of Human Society over time:

The documents on their vision say (citation from [Japanese Government, 2017])

"We aim at creating a society where we can resolve various social challenges by incorporating the innovations of the fourth industrial revolution (e.g. IoT, big data, artificial intelligence (AI), robot, and the sharing economy) into every industry and social life. By doing so the society of the future will be one in which new values and services are created continuously, making peoples' lives more conformable and sustainable.

This is Society 5.0, a super-smart society. Japan will take the lead to realize this ahead of the rest of the world."

The concept of "Society 5.0" was drafted in the 5th Science and Technology Basic Plan by the Council for Science, Technology and Innovation 2016.

Key issues of this plan are sustainability and social benefit for all citizen ("inclusion") by utilizing the advanced opportunities of Digital Transformation and Smart Technologies. Social reform (innovation) will achieve a forward-looking society that breaks the existing stagnation in societal and economic stagnation, forming a society of mutual respect, transcending the generations, and a active and enjoyable life for every person. This concept should also face challenges on a global scale, like depletion of natural resources, global warming, growing economic disparity, and even terrorism, by sharing knowledge and information, and cooperation. On the other hand, privacy and human independence is not such a key issue as in Europe, and sensibility is different from our society.

From a moral and ethical point of view, this sounds extremely well. The question that arises if it fits to other kinds of societies and regions in the world, having different social and economic preconditions, and different long-standing cultures and mind sets. From experience in history we know, that the fantastic technologies and opportunities can be misused by some people, criminals or governments against citizens. Therefore, in Europe, which is by far not so homogeneous in culture and mutual influences of neighbours as Japan, and America, institutions of different type have tried to set up ethical guidelines in the field of automated systems, cognitive decision taking (AI), and governance of data and connectivity.

6. Ethics Guidelines

6.1. Ethics Commission for Automated Driving (German Federal Ministry of Transport and Digital Infrastructure)

Automotive is a real mass market, and the trend towards highly automated and autonomous driving is ongoing in research and development. On the one hand, it is rather simple compared to complex AI issues and cognitive decision in achieving transparency in self-learning systems in detail (which is not possible at the moment), but even high-level principles are interesting.

The German document (available in English) [German Federal Ministry, 2017], includes a punctuation of 20 ethical rules for automated and connected vehicular traffic (shortened):

- The protection of individuals (human life) takes precedence over all other considerations (the systems must be programmed to accept damage to animals or property in a conflict).

- Design the vehicles such that they drive in a defensive and anticipatory manner, posing as little risk as possible to vulnerable road users (VRUs).

- In the event of unavoidable accident situations, any distinction based on personal features (age, gender, physical or mental constitution) is strictly prohibited. It is also prohibited to offset victims against one another.

- Design such that the need for an abrupt handover of control to the driver is virtually obviated. The systems must adapt more to human communicative behaviour rather than requiring humans to adapt.

- International standardization of the handover procedures.

- The **accountability** that was previously the sole preserve of the individual **shifts from the motorist to the manufacturers and operators** and to the **bodies responsible** for taking infrastructure, policy and legal decisions.

6.2. Trustworthy AI (Highly Automated systems in General)

Here is only a short overview on a few AI-related documents on "Trustworthy AI" and Ethical Guidelines for "Smart Systems' Decision Taking" provided:

- The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (AI/AS) (April 2016)
    - Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems (EAD V1 released)
    - Identification and recommendation of ideas for Standards Projects focused on prioritizing ethical considerations in AI/AS.
    - IEEE ECAIS "Ethics Certification for Autonomous and Intelligent Systems" (Industry Connections Activity Initiation Sept. 2018!).
- IEC/SMB Ad Hoc Group on autonomous systems and ethics (AHG 79) (ISO/TC299, June 20, 2018!!),
- "When Computers Decide" – European Recommendations on Machine Learned Automated Decision Making (Informatics Europe & EUACM 2018) includes Technical, Ethical, Legal, Economic, Societal and Educational recommendations),
- EC: "Ethics Guidelines for Trustworthy AI" (Final report April 2019, HLEG AI) https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

7. Conclusions

The technologically oriented funding organizations and the EC have a very positive approach and high expectations concerning the benefits of digitisation of economy, industry and society. The "Green Deal" programme and the human implications of these technologies, empowering people through education and skills, and on protecting ourselves with respect to the risks of these technologies, are targeting resilience and sustainability of society and economy.

The Japanese approach to "Society 5.0" even goes far beyond, and the visionary declarations are of high ethical and moral value. Applications like military, espionage etc. are explicitly excluded in research here. However, we should be aware that many of the achievements could be used against us as well (and some research projects consider this fact already) or lead to wrong decisions because of badly trained or biased AI systems. Knowledge and information can build a better society, but also be used against us by criminals or organizations. This requires careful European and international legislation and control to avoid the worst outcomes of these new technologies, and requires high public awareness. Politics sometimes tend to use safety and security threats as argument for more surveillance and control of people, endangering freedom and human rights. A first approach is taken by several authorities and international or governmental organisations to provide guidelines and recommendations for an ethical approach to highly autonomous systems.

The final question is: Will technologies of the fourth/fifth industrial revolution (e.g. IoT, Big Data, Artificial Intelligence, Connectivity, robots) enabling disruptive developments (evolutionary or revolutionary), lead to a resilient, sustainable and sharing society, to "well-being for all", if based on higher ethical values?

8. Acknowledgements

9. References

URL-Links were valid July 12, 2020.

European Commission, Independent High-Level Expert Group, "Ethics Guidelines for Trustworthy AI" (Final report April 2019, HLEG AI), Brussels; also available at https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Federal Ministry of Transport and Digital Infrastructure, Ethics Commission on "Automated and Connected Driving – Report June 2017", Germany; Summary available in English on https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?__blob=publicationFile

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems", First Edition. IEEE, 2019. Also available at https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

Informatics Europe and ACM Europe, "When computers decide", https://www.acm.org/binaries/content/assets/public-policy/ie-euacm-adm-report-2018.pdf

SSON – The Worlds largest shared Services and Outsourcing Network, "Reality Bites" (sponsored by edgeserve, Infosys BPM), https://www.ssonetwork.com/global-business-services/reports/covid-19-impact-on-global-service-delivery-models-reality-bites#form

CRISP Research, by Cloudflight 2020, ©, "Why European companies need to complete the digital transformation to remain viable and competitive"

AIOTI – Alliance for Internet of Things Innovation, http://www.aioti.org/resources/

ARTEMIS Strategic Research Agenda 2016, ARTEMIS Industrial Association, Eindhoven, NL.

EPoSS Strategic Research Agenda of the European Technology Platform on Smart Systems Integration, 2017. http://www.smart-systems-integration.org/public/documents/publications/EPoSS_SRA2017.pdf/view

European Commission, (2017a). White Paper on the Future of Europe, Brussels, European Commission (https://ec.europa.eu/commission/sites/beta-political/files/white_paper_on_the_future_of_europe_en.pdf)

European Commission, (2018). "Digitising European Industry – Two years after the launch of the initiative", Brochure March 2018, ISBN 978-92-79-80325-3, doi:10.2759/024187 https://ec.europa.eu/digital-single-market/en/news/digitising-european-industry-2-years-brochure

European Commission, (2019). Ethics Guidelines for Trustworthy AI, EC, High Level Expert Group on AI, April 2019, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

German Federal Ministry of Transport and digital Infrastructure, (2017). Ethics Commission on Automated and Connected Driving – Report June 2017 (Extract)

https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?__blob=publicationFile

Japanese Government, Cabinet Office (2018), Achieving Society 5.0, https://www8.cao.go.jp/cstp/english/society5_0/index.html

Japanese Government, (2017). Realizing Society 5.0. https://www.japan.go.jp/abenomics/_userdata/abenomics/pdf/society_5.0.pdf

Mayumi Fukuyama, (2018). Society 5.0: Aiming for a New Human-Centered Society, Japan Spotlight, July/August 2018, p. 47-50.

Prime Minister's Office of Japan, (2019). Society 5.0, Concept, https://www.youtube.com/watch?v=SYrv6kOsU1o Highlights, https://www.youtube.com/watch?v=S9JMuwvzz8g Human Ability, https://www.youtube.com/watch?v=odjuqbLJRMY

Schoitsch, E. (2019). "Beyond Smart Systems – Creating a Society of the Future (5.0): Resolving Disruptive Changes and Social Challenges", IDIMT 2019, Innovation and Transformation in a Digital World, Proceedings, Trauner Verlag, Linz, Austria, Schriftenreihe Informatik 48, p. 387-400 (ISBN 978-3-99062-590-3)

Schoitsch, E. (2018). "Smart Systems Everywhere – Intelligence, Autonomy, Technology and Society", IDIMT 2018, Proceedings, p. 153-165, Trauner Verlag, Reihe Informatik 47, 2018.

UNESCO (2019). Japan pushing ahead with Society 5.0 to overcome chronic social challenges, https://en.unesco.org/news/japan-pushing-ahead-society-50-overcome-chronic-social-challenges

Von der Leyen, U. (2019). A Union that strives for more – My agenda for Europe. https://www.europarl.europa.eu/resources/library/media/20190716RES57231/20190716RES57231.pdf

THREATGET: NEW APPROACH TOWARDS AUTOMOTIVE SECURITY-BY-DESIGN

Abdelkader Magdy Shaaban, Christoph Schmittner

Center for Digital Safety & Security
Austrian Institute of Technology, Vienna, Austria
abdelkader.shaaban@ait.ac.at, christoph.schmittner@ait.ac.at

Keywords

Potential Threats, Threat Modelling, Risk Assessment, and Automotive Domain

Abstract

Cybersecurity has become one of the biggest challenges in the automotive engineering process. Identifying the exact potential threats that can affect the vehicle system in the early stages of the vehicle life cycle is necessary because, once the vehicle is designed, it will be difficult to add security. This paper introduces ThreatGet as a novel threat modeling methodology, specifically developed for the automotive sector. ThreatGet can be integrated with the early stages of vehicle system design to identify threats and evaluate the associated risk to measure the overall risk and define security goals.

1. Motivational Background

Cyber-Physical Systems (CPS) are complex structures that incorporate logical and physical components with additional data sources such as cloud data services. The CPS plays a key role in investing significant resources in smarter products in the automotive sector. Vehicles themselves may be connected to other vehicles or infrastructure units (König et al., 2018). Like any connected device to the internet, modern vehicles are exposed to be attacked by different malicious ways. In addition to that, inside a modern vehicle is a hidden complexity of up to 150 Control Units connected via multiple communication networks, controlling a large number of sensors and actors with an increase in automated decision making. Therefore, cybersecurity requires to be a part of the developing phases in vehicular manufacturing (Shaaban et al., 2019). The World Forum on Vehicle Harmonization WP.29 provides a framework for harmonized vehicle regulation (UNECE, 2020b). The primary duty of the WP.29 is to keep the vehicle regulations up-to-date and appropriate, particularly in terms of technology, safety, and environmental change. The WP.29 forms the largest global vehicle regulatory framework, and it is a part of the United Nations Economic Commission for Europe (UNECE)[9] (ARGUS, 2020). In June 2018, WP.29 changed the Brakes and Running Gear (GRRF) Working Party into a new Automated/Autonomous and Connected Vehicles (GRVA) Working Party (UNECE, 2020a). This working party developed a new regulation that requires the car manufacturer to prove that security is adequately taken into account in the processes used under their cybersecurity management system (CSMS). Besides, it describes the baseline for threats, vulnerabilities, and attack methods in the automotive domain. It also addresses the security mitigation mechanisms against

---

9 https://www.unece.org/info/ece-homepage.html

threats intended to protect vehicles and outside of the vehicle (i.e., back-end server) from alternative cyber-attacks (Task Force, 2020).

In the future, the manufacturers of vehicles will need to check the cybersecurity of their vehicle systems before their products can be accepted for sale. Every three years, they have to demonstrate that they have implemented a certified cybersecurity management program for covering all processes, from vehicle engineering to documentation (AIT - Austrian Institute of Technology, LieberLieber Software, 2019). Furthermore, cybersecurity should be an essential part of the vehicular manufacturing process. Cybersecurity plays a key role in the vehicle domain because it protects critical vehicular units and software, which controls functional safety in a vehicle from various attack scenarios. Accordingly, the safety-security relationship is considered directly proportional as any malicious code inserted into any components or busses in a vehicle leads to damage or failure of either a particular unit or the whole vehicular system (Shaaban et al., 2019). New automotive systems will need to ensure safety and security; here, security analysis is considered the critical factor in this process (Ma & Schmittner, 2016). Threat modeling deals with this issue by defining an abstract model of threats, which is applied to a system to recognize potential threats that threaten the system (El Sadany et al., 2019). Threat modeling technique is an approach used for defining the security threats and vulnerabilities in the system model, and then determine the most cost-effective security measures that will be able to mitigate the threat (Abomhara et al., 2015). This work discusses a novel threat modelling methodology which is called ThreatGet. ThreatGet automates the threats analysis process and supports the security mitigation for reducing the overall risk in the vehicular domain.

2. Approach

ThreatGet is introduced to help in identifying potential threats in the system model in the early stages of the development process. It is an Enterprise Architect (EA) based tool is developed by the AIT Austrian Institute of Technology[10]. EA by Sparx Systems[11] is a commonly used platform for model-based systems engineering. ThreatGet aims to analyze the entities, artifacts, and links in the model design to identify the potential threats and risk assessment (El Sadany et al., 2019). It contains three main parts; the first part is the EA Add-in, the second and the third are web-based front-end and back-end sides. Figure 1 illustrates the main building blocks of ThreatGet.

2.1. EA Add-in

The ThreatGet Add-in can be easily integrated into the EA on the local user machine. The Add-in supports four separate phases in the threat modeling process, starting from modeling the system design until reporting the result into full documentation.

Modelling: is the phase, where the user collects all information related to the system structure and creates a diagram of the system model similar to system design, but focused on security. This model illustrates a realistic image of the interaction between the entities in the vehicular domain. Furthermore, this model is considered a living document containing the target changes and is kept up-to-date (Smith, 2016). ThreatGet supports the most common vehicular units and communication protocols that can be used to design a wide range of scenarios to define the internal interactions between the vehicular components. Plus, it supports different components that can be used to define

---

10 https://www.ait.ac.at/en/

11 https://www.sparxsystems.eu

the external interaction between the vehicle and the external environment, such as roadside units (RSUs).
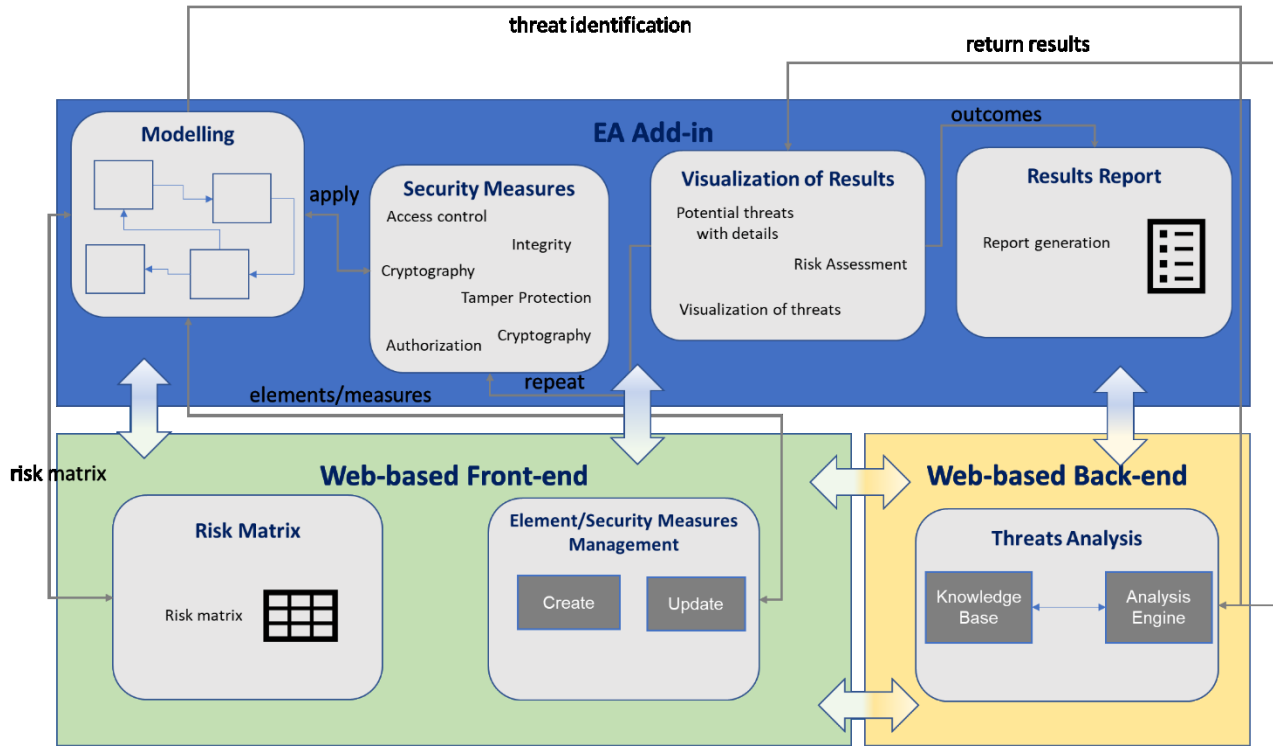
Figure 1 The building blocks of ThreatGet

Security Measures: play an integral role in the threat modeling process; they are used to examine how vehicle units and their interactions are secure against a wide range of potential threats. Furthermore, the user selects the applicable security properties that should be integrated into the vehicular units as a protective methodology for keeping the vehicle secure. These measures are used in ThreatGet to describe the security and analysis of related properties for elements (El Sadany et al., 2019). ThreatGet provides an exhaustive list of security measures to select the most suitable for the vehicle. The user selects a set of these measures to be a reflective image of the vehicle design; then, these measures are used in the threat analysis process to check how secure the vehicle model is and what are the security weaknesses that need more security concerns.

Visualization of Results: aims to display all outcomes (i.e., identified threats) and determine the exact level of risk per each threat detected. ThreatGet covers a wide range of threats, which are categorized according to the Microsoft STRIDE model (Shostack, 2014). Each threat is classified into one of the STRIDE categories according to its malicious activity. These categories are discussed in (Abomhara et al., 2015) as follows:

- **S**poofing: it means trying to get unauthorized access through false identity.

- **T**ampering: it intends to data modification by an unauthorized method.

- **R**epudiation: denying an action that is performed by legal/illegal user is called repudiation.

- **I**nformation Disclosure: it means private data has been revealed by an undesirable method.

- **D**enial of Service: it is an action to make a specific service, system, or application unavailable.

- **E**levation of Privilege: if a user with restricted authorization claims has a higher privilege, this action is called privilege elevation.

Afterward, ThreatGet performs a risk assessment process, which gives an overview of the overall risk. A risk assessment process is an essential approach for estimating and evaluating the risk severity level for each threat (Ramesh et al., 2017). This process is based on the impact and likelihood parameter values (Shaaban et al., 2019). ThreatGet supports different degrees of likelihood and impact. Users can handle these levels to determine the most effective risk assessment approach in different applications such as automotive, railways, Internet-of-Things (IoT), CPS, etc. ThreatGet defines five levels of risk (e.g., from 1 to 5); level 5 is the highest severity level (e.g., critical/extreme), where level 1 is the lowest risk (e.g., low). These levels explain the exact severity of the overall risk of the system model. The advantage of that, on the first hand, it is highlighting the security weaknesses in the system design, which need more security concerns. On the other hand, it allows to focus on the risks that could not be accepted by the user.

Results Report: after all potential threats in the system model are detected, the user can generate full documentation on all the details of the analysis process (i.e., threats, likelihood, impact, risk, and the units affected). This process helps collect all system model information as evidence of the system design's status and associated threats that could have a negative impact on system architecture.

2.2. Web-based Front-end (Management)

This is ThreatGet's second part as web interface; this interface gives users more flexibility in extending ThreatGet capabilities by handling components, security measures, and risk matrix. A web-based front-end is chosen here to support collaboration, e.g., multiple experts can work on the same knowledge base.

Element/Security Measures Management: ThreatGet offers a wide variety of vehicular components and security measures. However, some components could not be defined in ThreatGet, or a specific security measure is needed. In this case, the user needs to create the required entity (i.e., component or security measure) to use as required. Therefore, ThreatGet provides this service to allow the user to create new components and customized set of security measures to expand the capabilities of ThreatGet in order to be applied in multiple application domains. The user can handle these services through the front-end web user interface, and then the newly created entities are automatically integrated into the Add-in to use in the modeling process.

Risk Matrix: this is additional services provided by ThreatGet to handle the parameter values of the impact and likelihood, as discussed in Section 2.1. ThreatGet provides this function to allow the user to define customized risk levels to fit using in a wide range of system designs. The modified values will be automatically updated and considered in the risk assessment process to match the exact user needs in the risk assessment process.

2.3. Web-based Back-end (Threat Analysis)

This is the backbone of the ThreatGet, which manages the process of threats analysis. This process comes after the user completes the system model to identify possible system architecture threats. The entire model is uploaded to the "Analysis Engine" on the back-end side with all relevant details. ThreaGet applies a wide range of pre-defined rules to the system model for detecting potential threats. These rules are defined in a specific language developed by AIT and stored in the back-end knowledge base. Then the analysis engine uses these rules and all system design information (i.e., components, security measures, connections, etc.) to define the exact threats in the system model. The detected potential threats are then sent to the EA Add-in to display locally.

3. Discussion

Modern vehicles include a broad range of electronic and electrical systems to handle and control vehicle functions. Integrating internet technology with modern cars exposes these vehicle types to cyberattacks (Shaaban et al., 2019). Therefore, if any cyber attack takes place on a vehicle, we could expect that it affects a safe operation on the vehicle. Thus, to minimize these risks, it is essential to consider the potential threat in the vehicular design process. For this reason, ThreatGet is developed in order to be incorporated with the early stages of vehicle design and assess risk to mitigate it. If we assume a simple example as an Infotainment unit receives data from internal interfaces like USB and CD. The infotainment unit is directly connected to the vehicle 's main ECU; this ECU has communication channels with another ECU, which regulates vehicle speed and steering. Via the vehicle-to-anything V2X unit, the vehicle has direct interaction with external connection points like another car and RSUs. Figure 2 illustrates this scenario using ThreatGet.
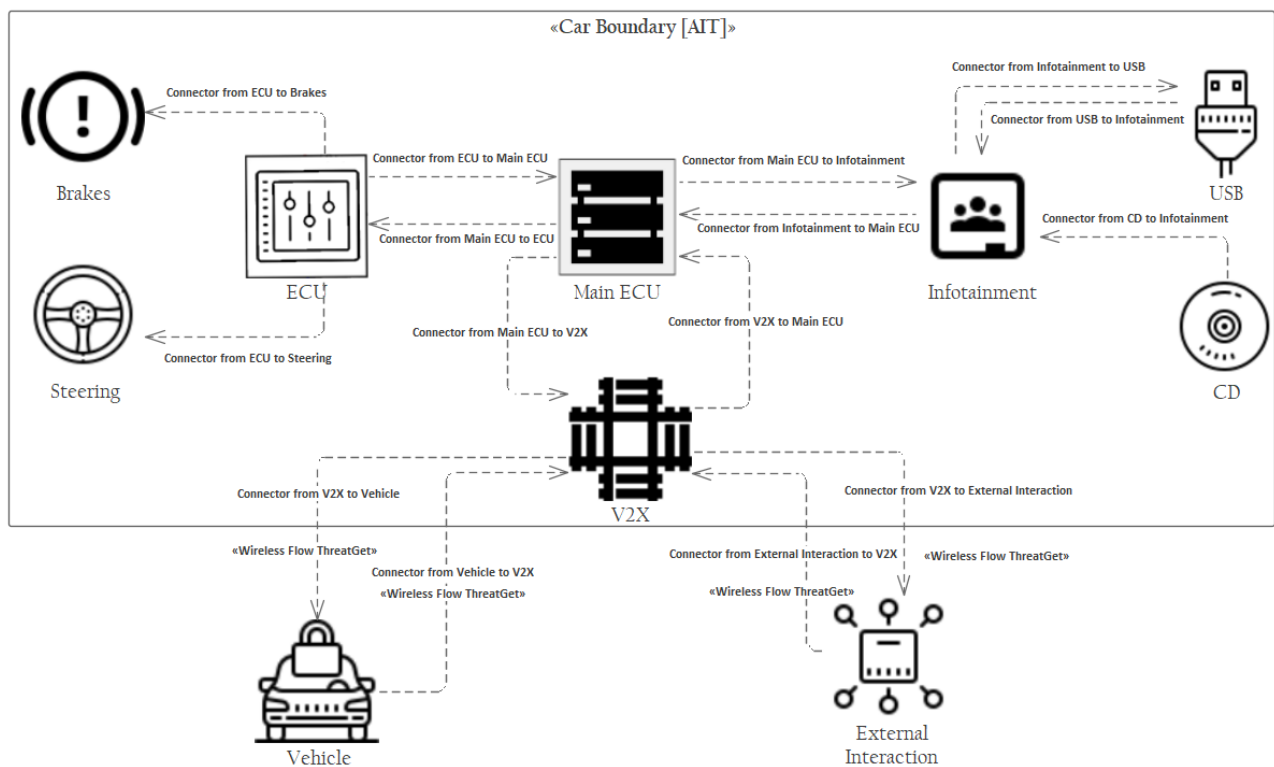
Figure 2 Internal and external interactions between different vehicular units

This example illustrates various communication channels between internal and external vehicle unit interactions. Attackers always try to exploit security vulnerabilities to initiate malicious actions against the vehicle system. For example, installing malicious software through the vehicle interface such as USB and CD could result in various potential consequences, such as the effect of safe vehicle operation, breach data confidentiality, or other negative consequences. ThreatGet helps to detect such a threat to easily mitigate it at the early stage of the vehicle life cycle. As shown in Figure 3, ThreatGet detects this threat, which is classified as "Tampering," according to the STRIDE category, with a risk level is five according to the likelihood and impact parameter values. For describing threats, ThreatGet provides a graphical summary of the elements affected by this threat. As illustrated by the figure, the USB and CD could be the way attackers can utilize to install malicious software to the Infotainment unit, which could affect the entire vehicular network.
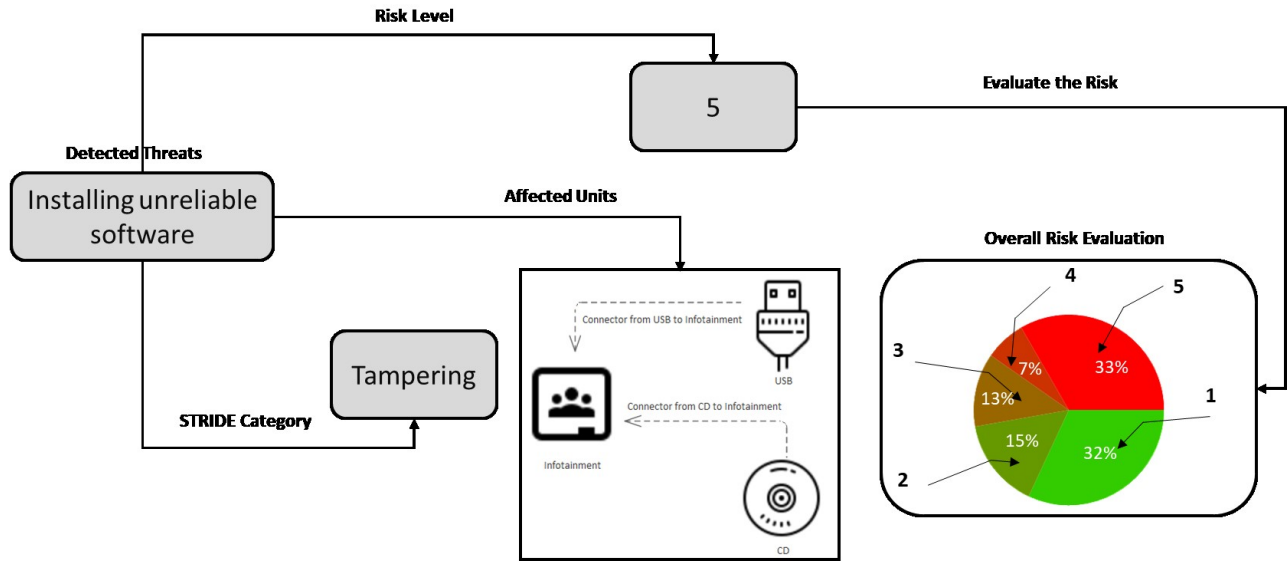
Figure 3 The detected threat and related information

ThreatGet detected other 71 threats; these threats are classified according to the STRIDE model. Table I illustrates the classification rates of all detected threats in this example.

Table 1 The rate of the identified threats according to the STRIDE model

Threat Category	Number of threats
Spoofing	5
Repudiation	11
Denial of Service	6
Tampering	37
Information Disclosure	6
Elevation of Privilege	7

Each threat has a specific risk level based on impact and likelihood values, so as illustrated in the pie-chart in Figure 3, ThreatGet assesses the overall risk of this example. The chart represents the overall risk evaluation by the tool for this diagram. For this diagram, ThreatGet assesses 33% of the threats classified as level five as the highest level of severity, where 32% of the threats are evaluated as level one of severity, indicating the lowest level of risk. The other evaluated risk levels (i.e., two, three, and four) are 15%, 13%, and 7% of the identified overall threats, respectively.

4. Summary and Conclusion

This paper introduces ThreatGet, a new effective threat modelling methodology to be integrated into several application domains (i.e., CPS and IoT), with the early stages of the system engineering process. The tool automatically performs a threat analysis to detect potential threats and support the risk management process in system design. In this study, an automotive example is discussed to illustrate a communication flow between alternative internal and external vehicular units. ThreatGet is applied to identify the expected threats in this simple vehicle design. ThreatGet performs a risk assessment for all threats individually, then assesses this model's overall risk. All ThreatGet results can be certified in full documentation reports. By this result, we can conclude that ThreatGet makes

security-by-design possible, not only in the automotive sector but also in other areas such as railways, energy, and critical infrastructure.

5. Acknowledgment

6. References

Abomhara, M., Gerdes, M., & Køien, G. M. (2015). A STRIDE-Based Threat Model for Telehealth Systems. 16.

AIT - Austrian Institute of Technology, LieberLieber Software. (2019). WHY THREATGET? https://www.ait.ac.at/fileadmin/mc/digital_safety_security/downloads/LieberLieber_Threatget_EN.pdf

ARGUS. (2020). UNECE WP29 Automotive Cybersecurity Regulation. Argus Cyber Security. https://argus-sec.com/unece-wp29-automotive-cybersecurity-regulation/

El Sadany, M., Schmittner, C., & Kastner, W. (2019). Assuring Compliance with Protection Profiles with ThreatGet. In A. Romanovsky, E. Troubitsyna, I. Gashi, E. Schoitsch, & F. Bitsch (Eds.), Computer Safety, Reliability, and Security (Vol. 11699, pp. 62–73). Springer International Publishing. https://doi.org/10.1007/978-3-030-26250-1_5

König, C. F. J., Meisl, G., Balcu, N., Vosseler, B., Hörmann, H., Höll, J., & Fäßler, V. (2018). Engineering of Cyber-Physical Systems in the Automotive Context: Case Study of a Range Prediction Assistant. In T. Margaria & B. Steffen (Eds.), Leveraging Applications of Formal Methods, Verification and Validation. Distributed Systems (Vol. 11246, pp. 461–476). Springer International Publishing. https://doi.org/10.1007/978-3-030-03424-5_31

Ma, Z., & Schmittner, C. (2016). Threat Modeling for Automotive Security Analysis. 333–339. https://doi.org/10.14257/astl.2016.139.68

Ramesh, R., Prabu, D. M., Magibalan, S., & Senthilkumar, P. (2017). Hazard Identification and Risk Assessment in Automotive Industry. International Journal of ChemTech Research, 7.

Shaaban, A. M., Schmittner, C., Gruber, T., Mohamed, A. B., Quirchmayr, G., & Schikuta, E. (2019). Ontology-Based Model for Automotive Security Verification and Validation. Proceedings of the 21st International Conference on Information Integration and Web-Based Applications & Services, 73–82. https://doi.org/10.1145/3366030.3366070

Shostack, A. (2014). Threat modeling: Designing for security. John Wiley & Sons.

Smith, C. (2016). The car hacker's handbook: A guide for the penetration tester. no starch press.

Task Force. (2020). Proposal for amendments to ECE/TRANS/WP.29/GRVA/2020/3 [Cyber Security and Over-the-Air issues].

UNECE. (2020a). Working Party on Automated/Autonomous and Connected Vehicles (GRVA). https://www.unece.org/trans/main/wp29/meeting_docs_grva.html

UNECE. (2020b). World Forum for Harmonization of Vehicle Regulation (WP.29). https://www.unece.org/trans/main/wp29/meeting_docs_wp29.html

EFFECTIVENESS AND SENSE OF PRIVACY POLICY STATEMENTS

Tomáš Sigmund

University of Economics, Prague
sigmund@vse.cz

Abstract

The paper researches and evaluates the effectiveness of privacy policy statements. These statements were the reaction on users' concerns concerning their privacy in the online environment at the beginning of the 21st century. However, it turned out that this solution based on transparency and free choice has been ignored by many users. The paper analyses possible reasons and finds similar trends in other areas of information society.

Keywords

privacy policy statement, privacy, information overload, confusion, ignoring privacy

1. Private information

In the area of web services, the potential for growth is very high because of Internet availability, low access costs, higher computer competence among users, availability of information, processing potential of ICTs, personalization of services and their convenience. On the other hand, users leave behind a lot of personal data which can be easily collected, compiled, used for manipulation or sold.

Culnan (2019) sees the big data as the current most threatening challenge for privacy policy as complex technologies, new data practices and ways of business operations and behavioural advertising are difficult if not impossible to explain in detail. Rapid changes make any privacy notice obsolete and inaccurate. Another problem represents the secondary use of data. Data may be used for purposes different from those originally stated, they may be combined with data from other sources, new personal data may be created from nonpersonal data. The understanding of contextual norms by companies is limited, even though they are important as they reflect expectations of acceptable practices in various contexts.

We should realize that the original equivalence model of companies' offer and users' choice is not valid any more. The asymmetry is big and users are very vulnerable and depend on companies how fairly they deal with their personal data. One option would be if companies integrated ethics and consumers' expectations into their risk assessment, but that is difficult to expect as a lot of data processing is done in secrecy, users very often don't know who and how their data uses and the benefits of personal data processing are numerous. Another problem represents the fact that ethics is not unanimous and transformable into algorithms. Trust can be defined with Jarvenpaa, Tractinsky & Saarinen (1999) as consumer's willingness to rely on the seller and to take actions in circumstances where such actions makes the consumer vulnerable to the seller. On the internet, users are vulnerable, as they have to rely on the service providers that they will use their private information appropriately. The internet is very tricky, the difference between reality and fiction is vague, users can be

manipulated or confused easily, they can't rely on their senses and social competencies as everything can be simulated and pretended.

Both the perception of privacy and trust is relative to the user and influences affecting him. The various models of consumer trust (Jarvenpaa, Tractinsky & Saarinen, 1999; Suh & Han, 2003; McKnight & Chervany, 2001) etc. came to the uniform conclusion that individual differences among consumers affect trust. These differences include awareness of internet risks and the past experience with it, inclination towards risk, gender, cultural background, trust in internet as a medium or trust in the service provider.

2. Privacy policy statement

To increase user trust web service providers can employ a lot of mechanisms that ensure the personal information will remain secure. Examples include approval of third party certifications, quality of the web site design, ratings, references, financial compensations for privacy breach or privacy policy statements. These statements codify how the personal information will be used. They are relatively simple and cheap and help especially in cases where the service provider doesn't have a high reputation yet. The privacy policy statements vary in length, ease of understanding, placement and level of protection Liu & Arnett (2002). It should be mentioned that the policies changed over the years significantly: 10 page fine print in legal jargon was changed to more user-friendly and user-understandably content. We will concentrate on the privacy policy statements to evaluate their effectiveness, problems and use them as paradigmatic examples for information society changes. The method of scientific literature research and analyses will be used and synthetic conclusions will be drawn.

As for privacy policy statements we need some universal rules that would somehow balance the situation even though they won't be able to respect individual differences. The internationally recognized and so most general principles of information privacy although less detailed than OECD or EU principles (the most famous is the GDPR) are the FIPP (Federal Trade Commission's Fair Information Practice Principles) practices. They consist of five dimensions: notice, choice, access, security, enforcement. The notice refers to information of data collection, choice to selection which personal information will be collected, access to possibility of accessing the stored information. Security means keeping accurate and secure data and enforcement concerns implementation and administration of policy by the company. These principles are the basics of the privacy statements which are then supplemented according to the specific circumstances and contexts.

2.1. Ignorance of privacy policy statements

The biggest problem with privacy policy statement is that consumers don't read them. As soon as in 2006 T. H. Cate in her book chapter (2006) pointed out that the amount of notices and consent opportunities rise, but the public ignores it. Some research has shown that consumers don't even read important information regarding the transaction like product warranties (Adler, 1994). Therefore, despite well written privacy policy statements they don't have any effect if users don't take notice of them. Agreeing to the terms and conditions is very often a fictitious answer. There are even ironic web pages (https://tosdr.org/ or @BiggestLie_com) pointing out the hypocrisy. The notice policy doesn't work. From one perspective it may seem that ignoring the privacy and terms of service policies is a regulatory failure and we need another maybe more pragmatic approach that would secure privacy and correct handling with data.

A recent study (Obar & Oeldorf-Hirsch, 2018) has shown that 74% of respondents skipped the privacy policy statements or terms of service statements. Most respondents agreed to the policies

(97% to privacy policy, 93% to terms of service). The negative predictor of reading was the information overload. Most users consider the policies a nuisance and ignore them. In spite of that, various strategies for the improvement of privacy policy statements are seen important to face the users concerns regarding privacy (FTC, 2012; Bakos, Marotta-Wurgler & Trossen, 2014).

Obar & Oeldorf-Hirsch (2018) in their study found out that people often ignore privacy and terms of service policies for social networks regardless if they sign up for a new service or if the policies change. The average time for reading the policies was 30 minutes, but the median reading time was 13.6 seconds, mean 73.72 seconds. The clickwrap allowing to bypass the policy and just to click to agree to it was used by 74% of respondents. That suggests that the implementation of clickwraps supports the ignoring behaviour. Even though the tested policies had problematic parts, 97% of participants agreed to privacy policy and 93% to the terms of service. A self-reported question on the behaviour regarding policies of big social networks shown that over 35% of respondents ignore them. Considering the privacy paradox the actual behaviour may differ significantly. Even the GDPR regulation does not provide a solution as it states in the section 32 that the subject's consent with processing his information may be just ticking a box and that the processes for gaining the consent must not be disruptive to the use of provided service.

Bakos, Marotta-Wurgler & Trossen (2014) carried out a research in 2007 on more than 48000 respondents and found out that terms of service policies were accessed by only 0.2% of them. The median time spent on the policy page was 30 seconds. A repeated similar research in 2012 shown that clickwraps have no impact on users. It should be noted on passing that the privacy paradox makes the investigation of policy behaviour more difficult as people usually say they take care of their privacy, but their actions are different. There is another reason why users ignore terms of service or privacy policies: They can't do anything about them anyway. If the customer wants to buy at a shop, he can either accept the conditions or he cannot buy anything.

Users' behavior regarding privacy differs. The type of information required by the web site can be classified as contact, biographical or financial (Meinert, Peterson, Criswell & Crossland, 2006). Their research has shown that users are more willing to provide contact than biographical information and least willing to provide financial information. Younger, educated and affluent users require a stronger protection than the older, less educated or poorer ones. The results however beside other things shows that 77.4% of the respondents had seen a privacy policy statement before the research, but only 45.6% had read it.

3. Research

We have carried out a research of users behavior in the online environment and their relationship towards privacy statements. The research used the method of questionnaire distributed online from 2nd March to 12th April 2020 through Google Forms. The respondents were both university students and general public. We received 108 answers to our questionnaire. 52% of our respondents were 15-25 years old, 26-37 years had 23% of our respondents and 25% were 38-69 years old. 37% of them were women, 63% were men. 60% of respondents were students, 6% worked in managerial positions, 6% worked in technical positions, 5% were retired, in services worked 5% of respondents, 4% were unemployed, 14% worked in other jobs. We used the 5-point Likert scale in the answers with 1 meaning definitely no and 5 definitely yes.

We asked our respondents three types of questions. The first category concerned understanding the situation and issues related to privacy. We asked these questions: Do you understand technologies processing big data and do you know what can be deduced from your personal information? (U1). Do you understand the general terms and conditions of various internet services? (U2) Do you understand the privacy statements of various internet services? (U3). Do you know what personal

data various services collect about you? (U4) Do you know where to find the privacy policy statements? (U5).

Table 1: Characteristics of questions concerning understanding privacy issues

Question	Mean	Standard deviation
U1	3,19	1,284
U2	2,78	1,105
U3	2,78	1,171
U4	3,06	1,229
U5	3,44	1,218
The Cronbach's alpha is 0,841.		

The second category concerned trust in the online environment in terms to privacy. The questions asked were: Do you trust e-shops in relation to your privacy? (T1) Do you trust social networks in relation to your privacy (T2) Do you trust internet search engines in relation to your privacy? (T3). Do you trust operating systems in relation to your privacy? (T4) Do you trust news portals in relation to your privacy? (T5) Do you trust discussion forums in relation to your privacy? (T6) Do you trust certificates confirming responsible handling with customers' data? (T7) Do you trust ratings and evaluations of internet services in relation to your personal data? (T8) Do you believe in the regulative role of financial fees I relation to privacy issues? (T9) Do you trust in the regulative role of law (GDPR etc.) in relation to privacy issues (T10)

Table 2: Characteristics of questions concerning trust

Question	Mean	Std. Deviation
T1	3,13	1,128
T2	1,94	0,895
T3	2,26	1,008
T4	3,06	1,214
T5	2,70	1,121
T6	2,31	1,020
T7	3,39	1,031
T8	2,98	1,050
T9	2,83	1,308
T10	3,28	1,259
The Cronbach's alpha is 0,803.		

The last category dealt with the behaviour in the online environment and approach towards privacy staements, terms and conditions of internet services. The questions asked were: Do you read terms and conditions of various internet services? (B1) Do you read the privacy statements of internet services? (B2) Do you automatically agree to privacy statements and policies of internet services? (B3).

Table 3: Characteristics of questions concerning behavior in privacy issues

Question	Mean	Std. Deviation
B1	2,50	1,219
B2	2,54	1,234
B3	2,53	1,233
The Cronbach's alpha is 0,794.		

The next step was preparation of a model by means of partial least square method. The Smart PLS software was used. The results show a moderate model fit (only the SRMR should be higher). The problem is the low extracted average variance at the Trust latent variable. We will try to improve it by excluding some variables with low loadings. The model characteristics can be seen below in tables 4 and 5. With bootstrapping the path coefficients can be seen in table 6.
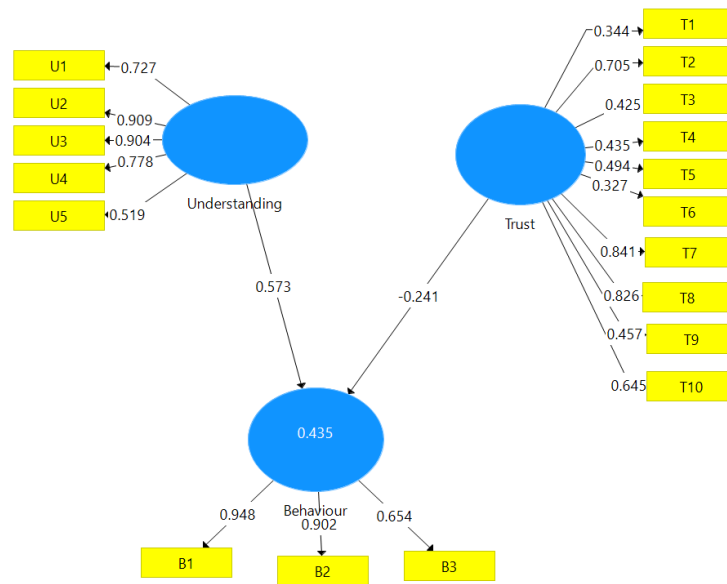
Figure 1: Model 1

Table 4: Fit measures of model 1

Fit measures	Saturated Model	Estimated Model
SRMR	0.138	0.138
d_ULS	3.262	3.262
d_G	1.287	1.287
Chi-Square	618.899	618.899
NFI	0.499	0.499

Table 5: Construct reliability and validity of model 1

Latent variable	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted
Behaviour	0.794	0.874	0.879	0.714
Trust	0.803	0.814	0.820	0.335
Understanding	0.841	0.945	0.883	0.609

Table 6: Path coefficients of model 1

Path	Original Sample	Sample Mean	Standard Deviation	T Statistics	P Values
Trust -> Behaviour	-0.241	-0.275	0.090	2.684	0.007
Understanding -> Behaviour	0.573	0.557	0.062	9.202	0.000

Table 4 shows the model fit – how well the model represents the data, table 5 the relevance of the latent variables cumulating the effects of individual variables and table 6 shows the coefficients of the latent variables in the regression model that affect the dependent variable behavior.

The outer loadings - they determine an item's absolute contribution to its assigned constructs - shown in the chart are in some cases not very high. That is why we have attempted to improve the model. After some improvements (we omitted U5, T5, T9 as their factor loadings were low, even though they were significant at the 5% significance level, and T1, T3, T4, T6 as they were not significant at the 5% significance level) we came to the following model. The model has improved to some extent. The total variance explained by Trust is higher, the model fit is better.
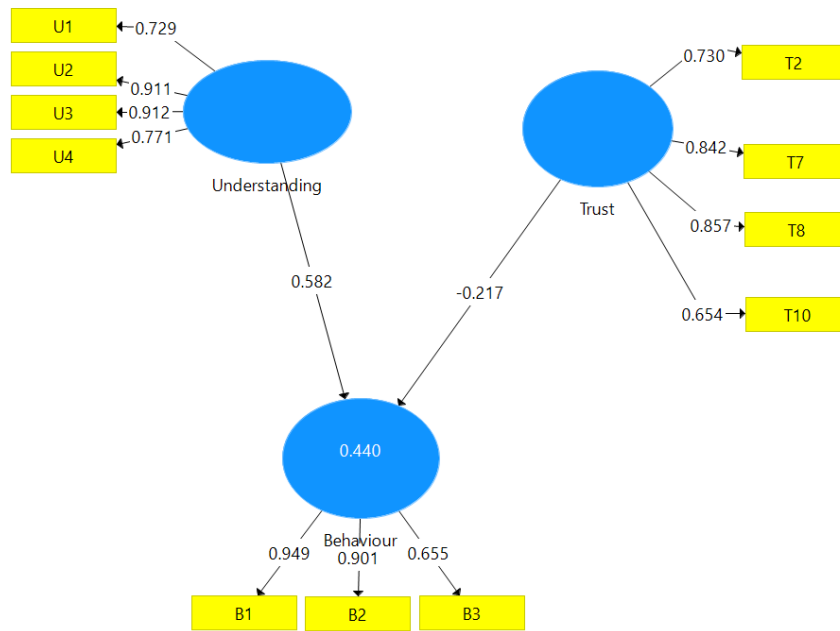
Figure 2: Model 2

Table 7: Fit measures of model 2

Fit measures	Saturated Model	Estimated Model
SRMR	0.128	0.128
d_ULS	1.079	1.079
d_G	0.414	0.414
Chi-Square	252.201	252.201
NFI	0.667	0.667

Table 8: Construct reliability and validity of model 2

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted
Behaviour	0.795	0.874	0.880	0.714
Trust	0.786	0.825	0.856	0.601
Understanding	0.859	0.929	0.901	0.696

Table 9: Path coefficients of model 2

Path	Original Sample	Sample Mean	Standard Deviation	T Statistics	P Values
Trust -> Behaviour	-0.217	-0.228	0.080	2.716	0.007
Understanding-> Behaviour	0.582	0.580	0.069	8.425	0.000

The model shows reading and considering terms and conditions of online services depends moderately on respondents understanding of technologies, the content of documents provided to users and their location. The better experience of understanding users have reading the terms and conditions of online services the more they read them. The model also shows weak negative dependence of reading and considering behavior on trust in internet services. The more users trust providers the less they read the documents on online services.

4. Reasons of privacy policy statements' ineffectiveness

In our research, the reading of privacy statement was not very high, the mean was in the middle of the Likert scale. The reasons of ineffective functioning of privacy notices and controls was analysed

by Schaub, Balabako & Cranor (2017). It may seem that users share personal data voluntarily as they extensively use many applications that collect their personal data, construct users' profiles and purchasing styles, the data are analysed, combined with other data and used for targeted advertising. The modern wearables are a recent invention that collects a lot of data. With the advent of internet of things privacy will be threatened even more. These devices are very popular and widely used.

However, the fact that many users take advantage of technologies that invade their privacy does not mean individuals surrender it, but that they feel helpless, lost and confused (Turow, Hennessy & Draper, 2015; Acquisti, Brandimarte & Loewenstein, 2015). The intrusions into privacy are often hidden and users don't know that their private data has been collected and processed, not to mention the fact that they can't predict how and when their information will be used. Users don't understand what data and how are used because such an understanding would require both technical and sociological expertise.

Schaub, Balabako's & Cranor's idea is to make privacy notices useful, usable and unobtrusive. They suggest the privacy notice to be 1) relevant in the current transactional context, 2) actionable, i.e. the user's choice and consent should be specific and explicit and 3) understandable, easy to use and not overloading with information. To realize these principles Schaub, Balabako & Cranor recommend 1) differentiating privacy policies according to various users (primary, secondary, incidental, protected); 2) provision of short and specific privacy notices adapted to specific context, system feature and audience; 3) highlighting unexpected and context violating practices with details on demand; 4) leveraging the notice by appropriate timing, channel (primary, secondary for e.g. fitness tracker or public in e.g. public spaces), modality (visual, auditory, haptic) and control (blocking, non-blocking, decoupled). Considering what has been said above it is questionable if these measures will secure effective privacy notices. And we mustn't forget the complexity and intricacy of today's IT world which does not allow for easy explanation of the practices and technical subtleties. Either the matter is too complicated or nobody can anticipate all details and directions where the data processing may lead. Privacy is an issue where the complexity and functioning of the world surpasses human ability to make a sense of it. The situation won't have a simple solution. A hint towards one would be if companies start to respect users, understand their expectations and take care that what they do is not only profit oriented, but also makes sense.

5. Conclusion

To conclude we may say that even though the privacy policy statements attempt to decrease the information asymmetry they collide with the users' inability to achieve the symmetry. A similar situation can be found in other examples of transparency which often lead to information overload. The world is becoming too complex and complicated and people are not able to understand all information necessary for their orientation and decision-making. The cybernetic law of requisite variety formulates it, too. But the problem as has been demonstrated on the example of privacy statements is not just excess of information, but also the working of technologies that is not translatable into the human categories. The technologies may use values that are not acceptable to humans and the processes they use in their decision-making (especially in the case of neural networks) can't be explained in human categories. The problem is currently acute in the area of privacy, but with the implementation and use of technologies is all areas of human life the consequences will be wider. The human world will so be assimilating to the absurd world of Franz Kafka's stories and its inhabitants will have similar feelings to the characters of existential novels.

6. Acknowledgement

7. References

Acquisti, A., Brandimarte, L. & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science. 347(6221). 509–514.

Adler, R. S. (1994). The last best argument for eliminating reliance from express warranties: "Real-world" consumers don't read warranties. South Carolina Law Review, 45(3), 429.

Bakos, Y., Marotta-Wurgler, F. & Trossen, D. R. (2014) Does anyone read the fine print? Consumer attention to standard-form contracts. The Journal of Legal Studies. 43(1), 1–35. doi.org/10.1086/674424

Cate, F. H. (2006). The failure of fair information practice principles. Consumer Protection in the Age of the Information Economy. J. K. Winn, rd., Surrey: Ashgate Publishing. 343–379.

Culnan, M. J. (2019). Policy to avoid a privacy disaster. Journal of the Association for Information Systems. 20(6). DOI: 10.17705/1.jais.00554

Federal Trade Commission. (2012). Protecting consumer privacy in an era of rapid change. Federal Trade Commission Report. Retrieved from https://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers

Jarvenpaa, S. L., Tractinsky, N. & Saarinen, L. (1999) Consumer trust in an Internet store: A cross-cultural validation. Retrieved from www.ascusc.org/jcmc/vol5/issues2/jarvenpaaa.html

Liu C. & Arnett, K. (2002). An examination of privacy policies in Fortune 500 Web sites. Mid-American Journal of Business, 17(1), 13-22.

McKnight, D. H. & Chervany, N. L. (2001). What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. International Journal of Electronic Commerce. 6(2), 35-59.

Meinert, D., Peterson, D., Criswell, J. & Crossland, M. (2006). Privacy policy statements and consumer willingness to provide personal information. JECO. 4(1). 1-17. DOI 10.4018/jeco.2006010101.

Obar, J. A. & Oeldorf-Hirsch, A. 2018. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. Information, Communication & Society. 23(1). DOI: 10.1080/1369118X.2018.1486870

Schaub, F., Balebako, R. & Cranor, L.F. (2017). Designing effective privacy notices and controls. IEEE Internet Computing. 21(1). 70-77.

Suh, B. & Han, I. (2003) The impact of customer trust and perception of security control on the acceptance of electronic commerce. International Journal of Electronic Commerce, 7(3), 135-161.

Turow, J., Hennessy, M. & Draper, N. (2015). The tradeoff fallacy: How marketers are misrepresenting american consumers and opening them up to exploitation. Annenberg School for Communication, University of Pennsylvania, Tech. Rep.

PEOPLE AND PROCESSES IN ENTERPRISE SOFTWARE DEVELOPMENT

ARE AGILE AND SCALED AGILE FRAMEWORKS REALLY ADDRESSING SOFTWARE DEVELOPMENT?

Alena Buchalcevova

Department of Information Technologies
Faculty of Informatics and Statistics
University of Economics, Prague
alena.buchalcevova@vse.cz

Keywords

Agile Development; Software development methods; Scaled Agile

Abstract

Digital transformation requires effective software development methods, techniques, and tools. Agile and Scaled Agile methods, that are broadly adopted in recent years, are believed to fulfil this goal. The aim of this paper is to analyze Agile and Scaled Agile methods and find out to what extent Agile and Scaled Agile methods cover software development processes. The analysis is performed through the mapping of individual methods to the ISO/IEC/IEEE 12207 international standard. The results of the mapping are presented in the form of graphical figures depicting all 12207 processes and indicating their level of adherence in color.

1. Introduction

In the recent years, digital transformation (DT) has emerged as an important phenomenon (Bharadwaj, El Sawy, Omar, Pavlou, Paul, & Venkatraman, 2013). Digital transformation is defined as "the use of new digital technologies (social media, mobile, analytics or embedded devices) to enable major business improvements (such as enhancing customer experience, streamlining operations or creating new business models)" (Fitzgerald, Kruschwitz, Bonnet, & Welch, 2013). Put differently, digital transformation is about adopting disruptive technologies to increase productivity, value creation, and social welfare (Ebert and Duarte 2018). As digital transformation impacts increasing complexity and scale of technological solutions and emphasizes time to market, quality, and affordability (Ebert and Duarte 2018), effective software development methods, techniques, and tools are needed to address these issues of IT systems delivery. Original Agile methods, designed to be used in small, single team projects (Boehm and Turner 2005), have stopped being sufficient. This has resulted in a birth of Scaled Agile methods that are nowadays broadly adopted according to global surveys (digital.ai, 2020; Dingsøyr & Moe, 2014).

Software development encompasses both a project management viewpoint and engineering viewpoint. Agile methods, especially the most popular ones such as Scrum and Kanban, cover mostly the management of software projects and do not address the engineering practices. From this viewpoint, is the case of Scaled Agile methods different? The aim of this paper is to analyze Agile and Scaled Agile methods and answer the research question: To what extent do Agile and Scaled Agile methods cover the engineering activities? To answer this research question, a mapping of individual methods to the ISO/IEC/IEEE 12207:2017 is provided.

The rest of the paper is organized as follows. Following the Introduction, Section 2 describes the background, i.e. a brief introduction to Agile and Scaled Agile methods and international standards for Software Process Improvement. Next, Section 3 describes the research approach. Section 4 then presents the results of the mapping of selected Agile and Scaled Agile methods to the ISO/IEC/IEEE 12207 standard. Finally, concluding remarks are given in Section 5.

2. Background

2.1. Agile and Scaled Agile Methods

While the traditional plan-driven software development methods, known as "Waterfall", do not scale to the challenges brought by digital transformation, the Agile and Lean approaches are a major step in that direction. Agile methods have now become the mainstream for software development worldwide. Formally, they were introduced through a set of four core values and 12 principles laid out in the Agile Manifesto (Beck et al., 2001). The benefits of Agile methods lie in reducing the risk of a project failure as software increments are delivered regularly based on prioritized requirements and frequent user feedback.

In order to use the Agile approach also for larger projects and larger companies handling an inter-team coordination and interfacing with other organizational units, such as human resources, marketing and sales, and product management is needed (Dikert, Paasivaara, & Lassenius, 2016). Therefore, a number of Scaled Agile methods and frameworks have emerged and are used in the industry such as the Discipline Agile Delivery (DAD), Large-scale Scrum (LeSS), Scaled Agile Framework (SAFe), Scrum@Scale, and Nexus (Kalenda, Hyna, & Rossi, 2018).

The state of Agile software development methods adoption has been surveyed both by scientists and practitioners since 2006. However, some surveys were focused only on specific geographic territories, e.g. Finland (Dikert et al., 2016) or Brazil (Campanelli, Camilo, & Parreiras, 2018). In addition, researchers have tried to reach English-speaking populations across the globe by offering them survey instruments in English (Kuhrmann et al., 2018; Kurapati, Manyam, & Petersen, 2012). In the world of business practice, the State of Agile™ survey with a global reach has been conducted by VersionOne and Agile Alliance annually since 2006. This well-known practitioner survey has also added a part focused on Scaled Agile methods since 2013. The recent (14th) edition (digital.ai, 2020) was carried out between August and December 2019. Based on this survey, the most used Agile methods for the analysis were selected, i.e. Scrum with a 58% usage, then the combination of Scrum and Extreme Programming with a 8% usage, quickly increasing Kanban with 7% and the combination of Scrum and Kanban with 10%. The Scaled Agile Framework (SAFe) with a 35% usage, Disciplined Agile Delivery (DAD) which an increasing usage reaching 4% and quite popular Large-scale Scrum (LeSS) framework with 4% were selected among Scaled Agile methods. These methods and frameworks are briefly introduced in next sections.

2.1.1. Scrum

Scrum is a development framework in which cross-functional teams develop products or projects in an iterative, incremental manner. It structures development in cycles of work called Sprints that are timeboxed. At the beginning of each Sprint, a cross-functional Team (of about seven people) selects items (customer requirements) from a prioritized list. The Team agrees on a collective target of what they believe they can deliver by the end of the Sprint. The Team gathers every day briefly to inspect its progress, and adjust the next steps needed to complete the work remaining. At the end of the Sprint, the Team reviews the Sprint with stakeholders, and demonstrates what has been built. This way, teammates obtain feedback that can be incorporated in the next Sprint. Scrum emphasizes a

working product at the end of the Sprint that is really "done"; in the case of software, this means a system that is integrated, fully tested, end-user documented, and potentially shippable (Deemer, Benefield, Larman, & Vodde, 2012).

2.1.2. Extreme Programming

Extreme Programming (XP) is a discipline of software development based on values of simplicity, communication, feedback, and courage (Beck, 1999). It works by using simple practices, with enough feedback to enable the team to see where they are and to tune the practices to their unique situation (Lindstrom & Jeffries, 2004). In XP, every contributor to the project is a member of the "Whole Team". The "Customer" is central to the team and works with them on daily basis. The team produces the software in a series of small, fully integrated releases that pass all the tests that the Customer has defined. Extreme Programmers work together in pairs and as a group, with a simple design and obsessively tested code, improving the design continually to keep it always just right for the current needs. The XP team keeps the system integrated and running all the time. The programmers write all production code in pairs, and all work together all the time. They code in a consistent style so that everyone can understand and improve all the code as needed (Lindstrom & Jeffries, 2004). The core XP practices are the Whole Team, Small Releases, Acceptance Tests, Pair Programming, Simple Design, Test-Driven Development, Refactoring, Continuous Integration, Team Code Ownership, and Coding Standards. Some of the XP practices such as testing, refactoring and continuous integration are heavily used within the Software Engineering discipline. However, the XP method stresses a synergy among its practices, which complete each other (Abrahamsson, Warsta, Siponen, & Ronkainen, 2003).

2.1.3. Kanban

Kanban is the most popular among Lean techniques which have their roots in Lean manufacturing and Toyota Production System. In 2004, Kanban was then firstly used for software development (Anderson, David, 2010). Kanban has five core principles: visualize workflow, limit work in progress, measure and manage flow, make process policies explicit, and use models to recognize improvement and opportunities (Anderson, David, 2010). The motivation to use Kanban was to visualize work, limit work in progress (WIP), and identify process constraints to achieve the flow and focus on a single item at a given time (Ahmad, Dennehy, Conboy, & Oivo, 2018). Various studies (Ahmad et al., 2018; Ahmad, Markkula, & Oivo, 2013; dos Santos, Beltrão, de Souza, & Travassos, 2018) have reported the benefits of using Kanban in software development, for example, a better visibility and understanding of the entire process, improved transparency of work and communication, better control of flow and WIP, improved team communication and coordination with other stakeholders, and increased customer satisfaction.

2.1.4. Scaled Agile Framework

The Scaled Agile Framework (SAFe) is a freely revealed knowledge base of proven, integrated patterns for enterprise-scale Lean-Agile development (Scaled Agile, 2018). The SAFe was created by Dean Leffingwell in 2012 and since then it has continually evolved to a current 5.0 version. The SAFe website (Scaled Agile, 2019) provides a guidance for scaling agile development across the Portfolio, Value Stream, Program, and Team levels that are part of the Big Picture, i.e. a visual overview of the Framework. The SAFe's practices are grounded on nine fundamental principles that have evolved from the Agile principles and methods, Lean product development, systems thinking, and observation of successful enterprises. The heart of the SAFe is the Program level, which revolves around an organization called the Agile Release Train (ART). Each ART aligns teams to a common

mission and vision via a single program backlog and produces valuable and evaluable system-level solutions every two weeks (Scaled Agile, 2018).

2.1.5. Disciplined Agile Delivery

The Disciplined Agile Delivery (DAD) framework is a hybrid of existing methods such as Scrum, Kanban, Agile Modelling, SAFe, Extreme Programming, Agile Data, Unified Process, and many others. DAD provides the flexibility to use various approaches and plugs the gaps not addressed by mainstream agile methods (Ambler & Lines, 2016). The main characteristics of this framework are that it: is a people first, learning oriented hybrid Agile/Lean approach; has a risk value delivery lifecycle; is goal-driven; is enterprise aware; is tactically scalable at the team level; and strategically scalable across all of the enterprise (PMI, 2019).

2.1.6. Large-scale Scrum

The Large-scale Scrum (LeSS) framework was created by Bas Vodde and Craig Larman in 2013 based on their experiences working with large-scale product development. As both authors state in (Larman & Vodde, 2016), scaling Scrum starts with understanding and being able to adopt standard one-team Scrum. Large-scale Scrum requires examining the purpose of single-team Scrum elements and figuring out how to reach the same purpose while staying within the constraints of the standard Scrum rules. LeSS provides two different large-scale Scrum frameworks, i.e. the basic LeSS applicable up to eight teams (of eight people each) and the LeSS Huge that introduces additional scaling elements for development up to hundreds of developers.

2.2. International Standards for Software Process Improvement

The traditional way towards more successful software projects is represented by a Software Process Improvement (SPI) initiative. Software Process Improvement aims to improve software processes (Humphrey, Watts, 1989) and comprises a variety of tasks, such as scoping, assessment, design and realization, and continuous improvement, etc. In this field, a number of SPI models have been developed like the Capability Maturity Model Integration (CMMI) (CMMI Institute, 2019) or international standards ISO/IEC/IEEE 12207 (ISO/IEC/IEEE, 2017) or ISO/IEC 15504 series that has been converting into the revised content associated with the ISO/IEC 33000 series. For small companies is then intended the ISO/IEC 29110 series (SC7, 2016). In 2017, a new version of the ISO/IEC/IEEE 12207 standard was published (ISO/IEC/IEEE, 2017) which establishes a common framework for software life cycle processes. It contains processes, activities, and tasks that are to be applied during the acquisition of a software system, product, or service and during the supply, development, operation, maintenance, and disposal of software products. It serves as the Process Reference Model for software life cycle processes.

3. Research Method

To answer the research question and identify the coverage of the engineering activities among Agile and Scaled Agile methods, the analysis of its extent among selected Agile and Scaled Agile methods is performed. The ISO/IEC/IEEE 12207:2017 processes were used as the basis for the analysis. Selected methods for the analysis, as stated in Section 2.1, were Scrum, Extreme programming (XP)(Beck, 1999), Kanban, Scaled Agile Framework (SAFe), Disciplined Agile Delivery (DAD) (Scaled Agile, 2018) and Large-scale Scrum (LeSS). The analysis was conducted in the form of the mapping of individual methods, factually their specific practices, onto the ISO/IEC/IEEE 12207:2017 processes. The level of coverage was expressed by the scale of (i) not covered; (ii)

partially covered; (iii) largely covered; (iv) fully covered. The adherence to the ISO/IEC/IEEE 12207:2017 processes was assessed based on the process outcomes and activities and tasks. The results of the analysis are presented in the form of graphical figures depicting all 12207 processes and indicating their level of adherence by color.

4. Mapping of Selected Methods to ISO/IEC/IEEE 12207 Standard

Figure 1 depicts the mapping of the Scrum method to the ISO/IEC/IEEE 12207:2017 processes. It is apparent that only technical management processes are covered by Scrum. Among technical processes, only Requirements definition processes are covered. As Scrum is oriented only on one project, no processes at the organizational level are covered. These results support the well-known classification of the Scrum method being a project management framework. In this sense, Scrum does not definitely cover software development sufficiently and other engineering methods and technics are needed to be combined with the Scrum method.
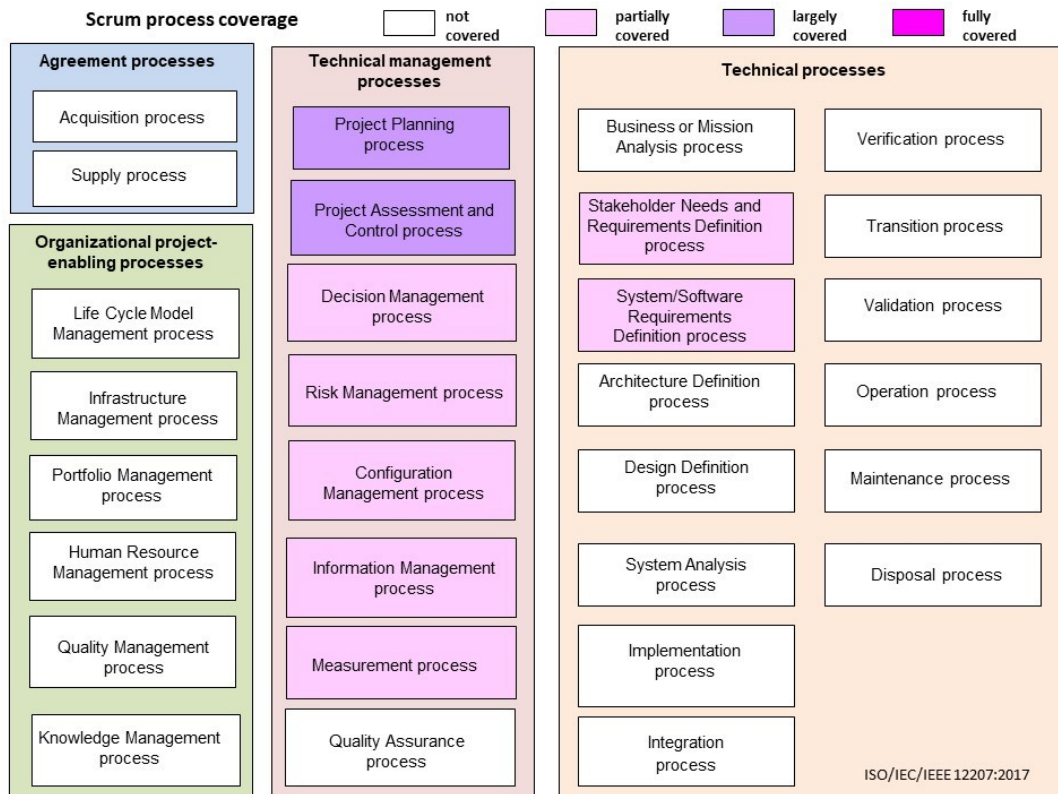
Figure 1. Mapping of the Scrum method to the ISO/IEC/IEEE 12207:2017

Figure 2 shows mapping of the Extreme Programming (XP) method to the ISO/IEC/IEEE 12207:2017 processes.
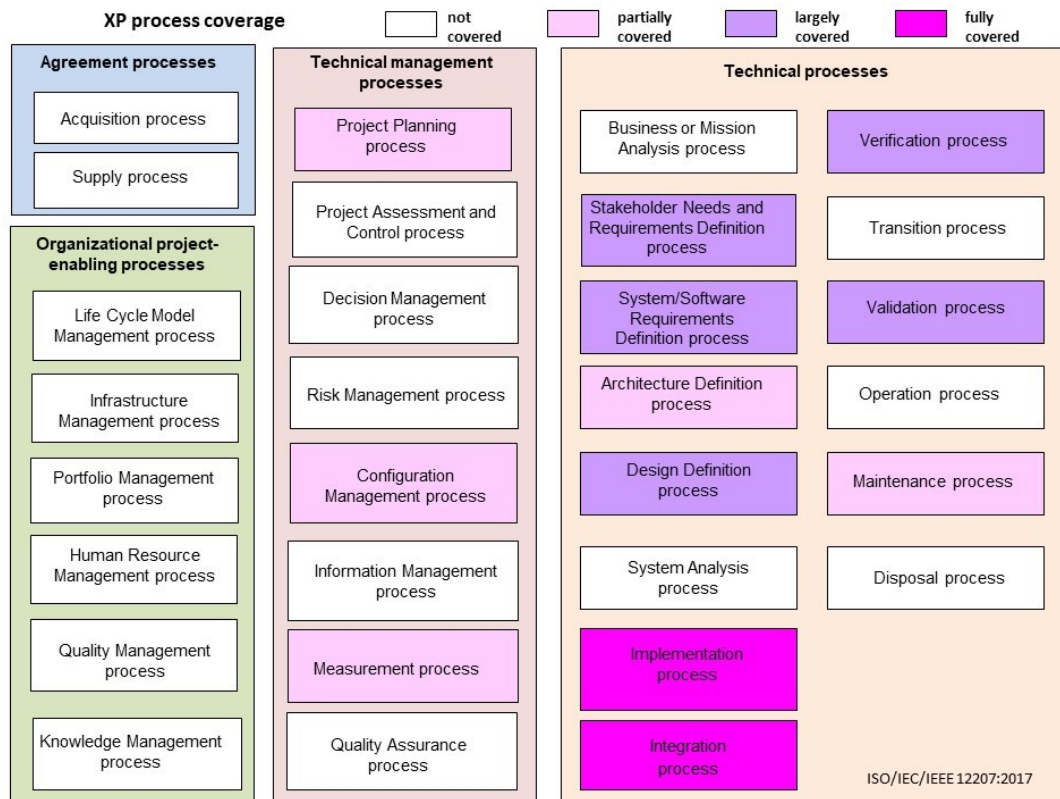
Figure 2. Mapping of the XP method to the ISO/IEC/IEEE 12207:2017
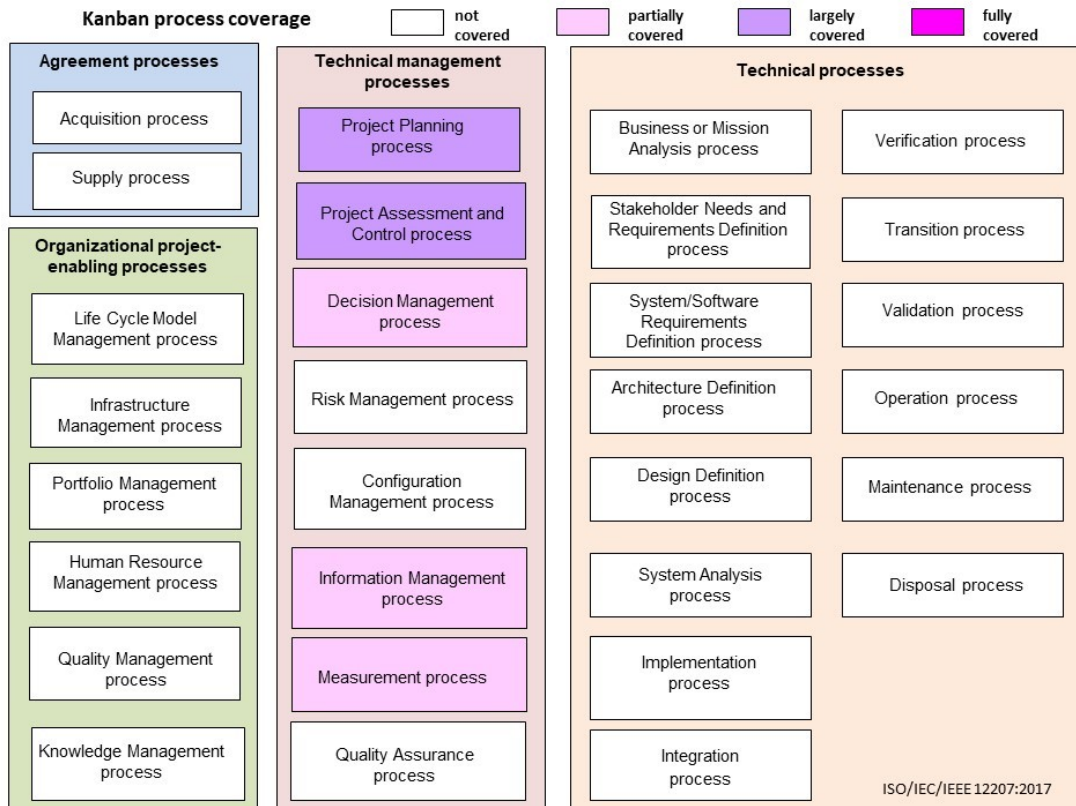
Figure 3. Mapping of the Kanban method to the ISO/IEC/IEEE 12207:2017

The Extreme Programming has been selected for the analysis as it is one of a few Agile methods that is specifically focused on software engineering. As seen from Figure 2, the results of the mapping confirm this statement. XP is a very disciplined approach based on the engineering practices such as Pair Programming, Simple Design, Test-Driven Development, Refactoring, Continuous Integration, Acceptance Tests and Coding Standards. On the other hand, XP does not cover the management processes, except for a partial coverage of the Project Planning, Configuration Management, and Measurement processes. This is why XP is combined with Scrum from its very beginning and this combined method still keeps a relatively high share (8%) among methods being used for software development (digital.ai, 2020).

Recently, the Kanban method has been gaining an increasing popularity. It is used either separately (7%) or as a combination with Scrum which is named Scrumban and is used by 10% of the latest State of Agile survey respondents (digital.ai, 2020). Figure 3 depicts the Kanban´s mapping to the ISO/IEC/IEEE 12207:2017 processes. Similar as Scrum, Kanban is focused especially on project management and is even more lightweight compared to Scrum. It defines only 5 principles that lead the process. No engineering practices are defined within Kanban, not even for requirements definition. The quite popular combination of Scrum and Kanban, i.e. Scrumban, however does not improve the coverage extent of the engineering practices. Also, in the case of Kanban and Scrumban, other engineering practices must be incorporated to support software development.

Figures 4 to 6 show the mapping of selected Scaled Agile methods to the ISO/IEC/IEEE 12207:2017 processes.
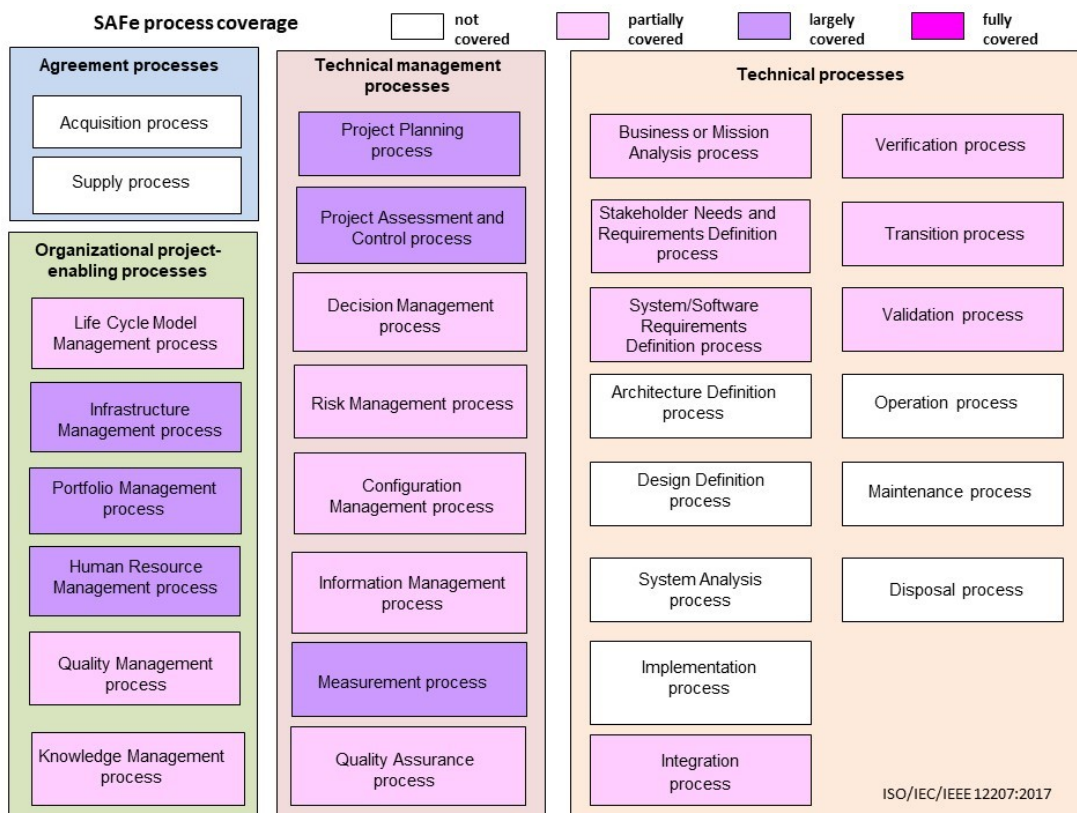
Figure 4. Mapping of the SAFe method to the ISO/IEC/IEEE 12207:2017
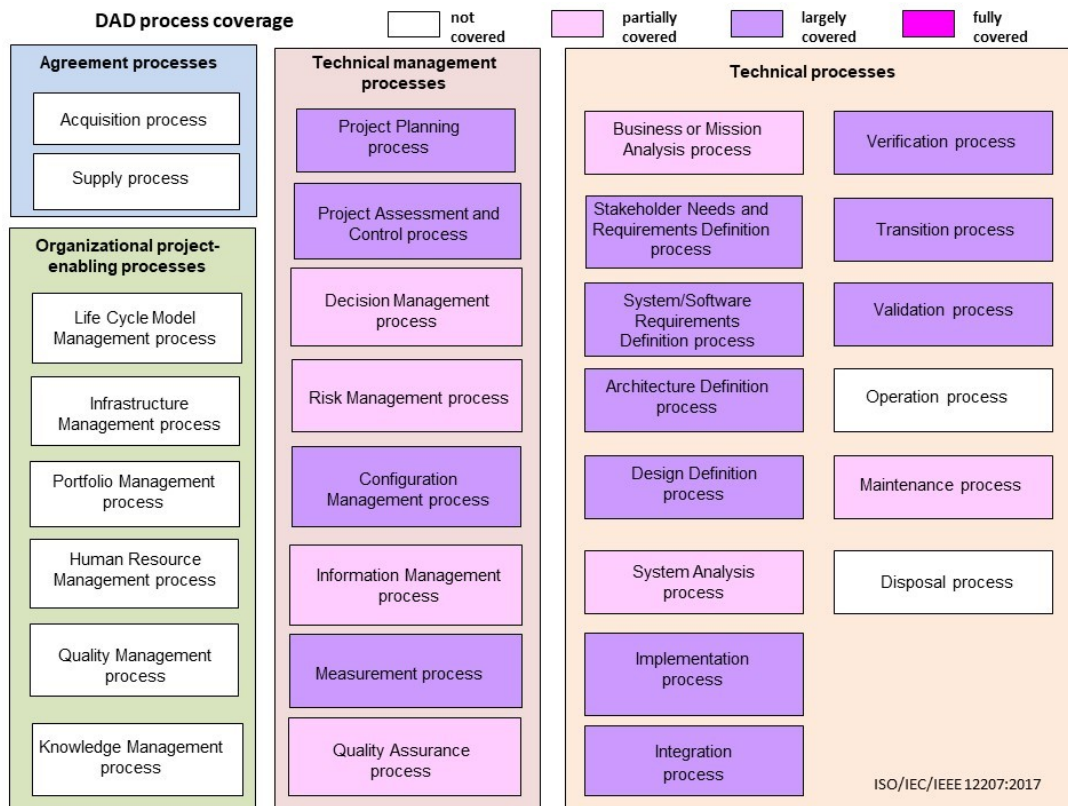
Figure 5. Mapping of the DAD method to the ISO/IEC/IEEE 12207:2017
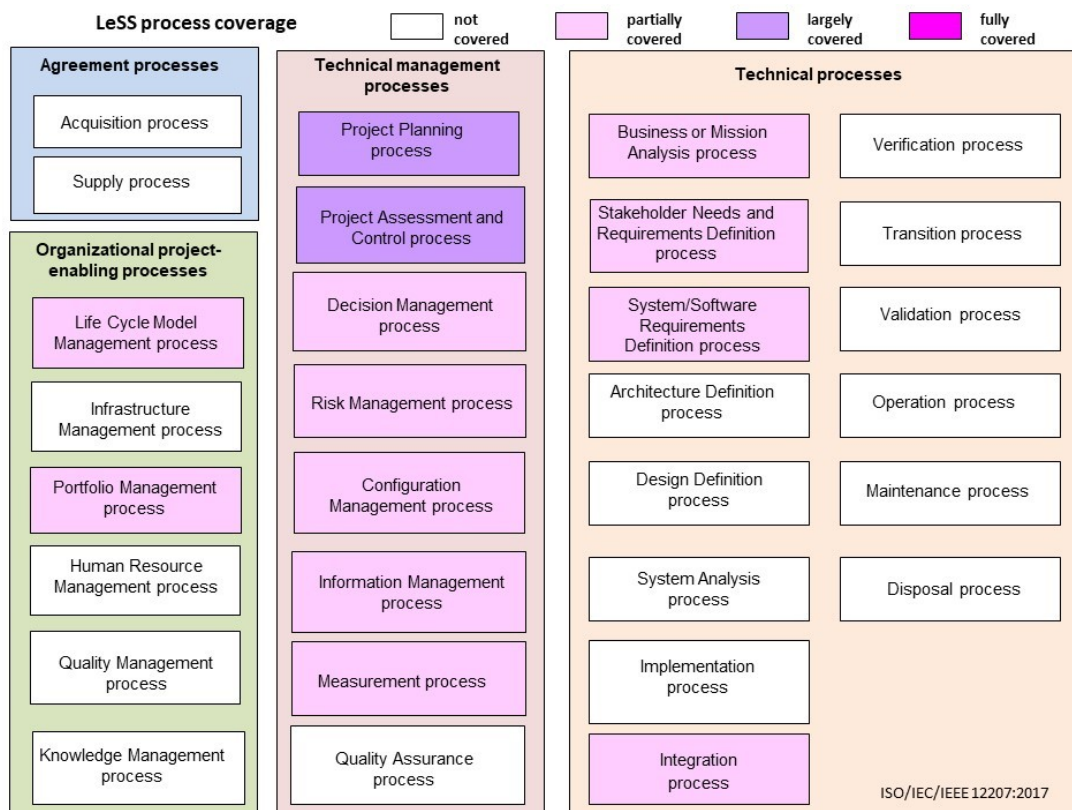
Figure 6. Mapping of the LeSS method to the ISO/IEC/IEEE 12207:2017

438

The mapping of the Scaled Agile Framework (SAFe), which is the mostly used Scaled Agile method nowadays, is depicted in Figure 4. Compared to other methods, much more 12207 processes are covered by the SAFe. All the technical management processes are covered largely or partially. Worth noticing is the coverage of the Organizational project enabling processes. The SAFe as an example of Scaled Agile methods is focused on the enterprise level and covers largely the Infrastructure Management, Portfolio Management and Human Resource Management processes. Other Organizational project enabling processes are covered partially. However, the SAFe does not cover some engineering processes like the Architecture Definition, Design Definition, System Analysis, and Implementation.

The Disciplined Agile Delivery (DAD) framework is the only method/framework which covers both technical management processes and technical processes as apparent from Figure 5. In this case, the processes are covered mostly largely. This high level of coverage is due to an effective combination of various proven software development methods like Rational Unified Process, Extreme Programming, Agile modelling, Scrum, Lean and others. This broad coverage of most software development processes may have been the cause of a growing popularity of this framework recently. Disciplined Agile Delivery is a part of the whole Disciplined Agile toolkit that encompasses four levels: Disciplined Agile Delivery, Disciplined DevOps, Disciplined Agile IT and Disciplined Agile Enterprise. Even though DAD does not cover the Organizational project-enabling processes, other parts of the Disciplined Agile toolkit do.

Figure 6 shows process coverage of the Large-scale Scrum (LeSS) framework. On the contrary to SAFe, LeSS covers much less Organizational project enabling processes. Technical management processes are covered mostly, but only partially. Quality assurance process is not covered. Similar to SAFe key technical processes are not covered in LeSS, except for requirements management and integration processes.

5. Conclusion

In this paper, the analysis of selected Agile and Scaled Agile methods was presented determining to what extent individual methods cover software development processes. The most used Agile and Scaled Agile methods were selected, i.e. Scrum, Extreme Programming, Kanban, Scaled Agile Framework (SAFe), Disciplined Agile Delivery (DAD) and Large-scale Scrum (LeSS). The analysis was performed through the mapping of individual methods to the processes of the ISO/IEC/IEEE 12207:2017 international standard.

The results of the mapping show that Scrum covers only the technical management processes and thus other engineering methods and technics must be used together with Scrum to support software development completely. On the opposite side lies Extreme Programming which is specifically focused on software engineering and does not cover project management sufficiently. That is why XP is usually combined with Scrum and this combined method keeps still a relatively high share in method usage. Kanban and Scrumban are similarly to Scrum mainly focused on the project management practices.

The Disciplined Agile Delivery (DAD) framework is the only method which covers both technical management processes and technical processes. It is a real hybrid framework that effectively combines several methods like Scrum, Kanban, Agile Modelling, Extreme Programming, Unified Process, and others.

The SAFe and LeSS are popular Scaled Agile methods and as such they cover the Organizational project enabling processes, i.e. the processes at an enterprise level. However, they do not cover key

engineering processes like the Architecture Definition, Design Definition, System Analysis, and Implementation.

Thus, to answer the research question, we can summarize that the most popular and most used Agile and Scaled Agile methods such as Scrum and SAFe do not cover software engineering processes sufficiently and have to be supplemented by other engineering practices.

6. Acknowledgement

7. References

Abrahamsson, P., Warsta, J., Siponen, M. T., & Ronkainen, J. (2003). New directions on agile methods: A comparative analysis. Proceedings - International Conference on Software Engineering, 6, 244–254. https://doi.org/10.1109/icse.2003.1201204

Ahmad, M. O., Dennehy, D., Conboy, K., & Oivo, M. (2018). Kanban in software engineering: A systematic mapping study. Journal of Systems and Software, 137, 96–113. https://doi.org/10.1016/j.jss.2017.11.045

Ahmad, M. O., Lwakatare, L. E., Kuvaja, P., Oivo, M., & Markkula, J. (2017). An empirical study of portfolio management and Kanban in agile and lean software companies. Journal of Software: Evolution and Process, 29(6). https://doi.org/10.1002/smr.1834

Ahmad, M. O., Markkula, J., & Oivo, M. (2013). Kanban in software development: A systematic literature review. Proceedings - 39th Euromicro Conference Series on Software Engineering and Advanced Applications, SEAA 2013, 9–16. https://doi.org/10.1109/SEAA.2013.28

Ambler, S. W., & Lines, M. (2016). The Disciplined Agile Process Decision Framework. https://doi.org/10.1007/978-3-319-27033-3_1

Anderson, David, J. (2010). Kanban: Successful Evolutionary Change for Your Technology Business. Blue Hole Press.

Beck, K. (1999). Extreme Programming Explained : Embrace Change. Addison-Wesley Professional.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., … Thomas, D. (2001). Manifesto for Agile Software Development. Retrieved March 10, 2019, from https://agilemanifesto.org/

Bharadwaj, A., El Sawy, Omar, A., Pavlou, Paul, A., & Venkatraman, N. (2013). Digital Business Strategy: Toward a Next Generation of Insights. MIS Quarterly, 37(20), 471–482.

Campanelli, A. S., Camilo, R. D., & Parreiras, F. S. (2018). The impact of tailoring criteria on agile practices adoption: A survey with novice agile practitioners in Brazil. Journal of Systems and Software, 137. https://doi.org/10.1016/j.jss.2017.12.012

CMMI Institute. (2019). CMMI. Retrieved August 10, 2019, from https://cmmiinstitute.com/

Deemer, P., Benefield, G., Larman, C., & Vodde, B. (2012). A Lightweight Guide to the Theory and Practice of Scrum.

digital.ai. (2020). 14th annual STATE OF AGILE Report. Retrieved from https://stateofagile.com/#ufh-c-7027494-state-of-agile

Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations : A systematic literature review. The Journal of Systems & Software, 119, 87–108.

Dingsøyr, T., & Moe, N. B. (2014). Towards Principles of Large-Scale Agile Development. International Conference on Agile Software Development. Springer, Cham.

dos Santos, P. S. M., Beltrão, A. C., de Souza, B. P., & Travassos, G. H. (2018). On the benefits and challenges of using kanban in software engineering: a structured synthesis study. Journal of Software Engineering Research and Development, 6(1), 1–30. https://doi.org/10.1186/s40411-018-0057-1

Fitzgerald, M., Kruschwitz, N., Bonnet, D., & Welch, M. (2013). Embracing Digital Technology: A New Strategic Imperative | Capgemini Consulting Worldwide. In MIT Sloan Management Review (Vol. 55). Retrieved from https://www.capgemini-consulting.com/SMR

Humphrey, Watts, S. (1989). Managing the software process. Boston: Addison-Wesley.

ISO/IEC/IEEE. (2017). ISO/IEC 12207:2017 Systems and Software Engineering -- Software Life Cycle Processes.

Kalenda, M., Hyna, P., & Rossi, B. (2018). Scaling agile in large organizations: Practices, challenges, and success factors. Journal of Software: Evolution and Process, 30(10), e1954.

Kuhrmann, M., Diebold, P., Munch, J., Tell, P., Trektere, K., Mc Caffery, F., … Prause, C. (2018). Hybrid Software Development Approaches in Practice: A European Perspective. IEEE Software. https://doi.org/10.1109/MS.2018.110161245

Kurapati, N., Manyam, V. S. C., & Petersen, K. (2012). Agile software development practice adoption survey. International Conference on Agile Software Development. Springer, Berlin, Heidelberg.

Larman, C., & Vodde, B. (2016). Large-Scale Scrum: More with LeSS. Addison-Wesley Professional.

Lindstrom, L., & Jeffries, R. (2004). Extreme programming and agile software development methodologies. Information Systems Management, 21(3), 41–52. https://doi.org/10.1201/1078/44432.21.3.20040601/82476.7

PMI. (2019). Disciplined Agile. Retrieved March 10, 2019, from https://www.disciplinedagiledelivery.com/

SC7. (2016). ISO/IEC TR 29110-1 Systems and software engineering -- Lifecycle profiles for Very Small Entities (VSEs) -- Part 1: Overview.

Scaled Agile. (2018). SAFe 4.6 Introduction. https://doi.org/10.1017/S1431927615014221

Scaled Agile. (2019). Scaled Agile Framework. Retrieved from https://www.scaledagileframework.com/

HAS PRODUCT OWNER'S ROLE CHANGED WITH AGILE DEVELOPMENT EXPANSION?

Daniel Remta

Department of Information Technologies
Faculty of Informatics and Statistics
University of Economics, Prague
xremd03@vse.cz

Keywords

Product Owner, Activities, Agile, Scrum, Software Development

Abstract

Agile as an approach to software development is spreading in industry and is being adopted by large organizations. The most used agile method is Scrum, which defines three roles: Product Owner, Scrum Master, and Development Team. The mentioned roles are undergoing changes with increased adoption rates. This paper focuses on the Product Owner role. Results from the literature review show that role has changed from a relatively simple role to the role covering a wide range of activities. Paper describes changes in PO activities since its introduction to current adoptions in large organizations.

1. Introduction

Agile as an approach to software development has become a common practice. 73% of organizations are using agile practices at least sometimes (Project management institute, 2018). Agile is incremental and iterative software development done by closely collaborating teams (Shammi et al., 2011). Agile as a methodology first gained attraction when Sutherland and Schwaber (1995) discussed the first agile process for software development called Scrum in the 1995 OOPSLA conference. Scrum became very popular, even though sometimes in a different form than its originators have had in mind (Dolezel et al., 2019), and is the most common agile method used (CollabNet VersionOne, 2019).

The latest Scrum guide (Schwaber & Sutherland, 2017) defines three roles: Product Owner (PO), Development Team, and Scrum Master (SM). In Scrum, the work is done in 1 to 4 weeks iterations called Sprints, each aiming to end up with a potentially releasable product increment (Schwaber & Sutherland, 2017). Every increment contains implemented requirements selected by the PO. The list of requirements is called a Product Backlog. The Development Team is self-organizing and works independently during the Sprints, ending each sprint with a demonstration of a potentially releasable increment. The Scrum Master's role is to eliminate any emerging impediment and enforce the Scrum Process (Eloranta et al., 2016).

Despite the original orientation of agile methods on small teams, they became popular also in large companies (Dyba & Dingsoyr, 2009; Paasivaara & Lassenius, 2011) and are there largely adopted and scaled (Dikert et al., 2016). Agile methods, in their original design, have stopped being sufficient (Buchalcevova, 2020). Tailoring agile methods to the particular requirements of organizations is an

absolute necessity (Lindvall et al., 2004). Hence instead of rigidly following the single method prescriptions, selecting, adapting, and combining software practices is a reality (Buchalcevova & Doležel, 2019). Considering the mentioned increase in adoption, scaling, and tailoring, an assumption that the activities covered by roles tight with Agile had to change was made. This paper aims to map and discuss the evolution of the PO role activities since Agile has started to spread in the industry. The PO role has been selected due to its identified criticality in the success of projects (Dikert et al., 2016), to help enterprises and research community understand changing requierements for the people positioned into the role. Mapping of the PO role activities through the lens of the time provides a view on the contemporary concept of the role, reflecting increasing role adoptions to the different kinds and sizes of organizations.

The main research question is: What changes in Product Owner's activities have occurred since the introduction of Scrum till the current role adoptions? Activity is understood as the work of a person, group, or organization to achieve something (Cambridge University Press, 2008).

To reach this goal, the following objectives (OBJn) were defined: (OBJ1) Identify publications mentioning PO activities, (OBJ2) Extract and categorize the activities, (OBJ3) Create the timeline representing the evolution of activities. The main contribution of this paper is to provide a view of the evolution and changes of PO activities over time.

The content of this paper is organized as follows. Section 2 describes the research method. In Section 3, the PO role is introduced, and the listing, chronological ordering, and placing of activities into categories are done. Then the timeline with a brief discussion is presented in section 4. Next, the validation of results is performed. The paper ends with a conclusion of the findings and recommendations for further research.

2. Research Method

To answer the research question, a literature review was conducted, activities extracted, categorized, and positioned on the timeline. Then, the findings were discussed and validated. In this section, the process steps are described in relation to study objectives. The last paragraph is dedicated to the validation step.

OBJ1 - Identify publications mentioning PO activities. The literature review was conducted using keywords: Scrum, Agile, Product Owner, Product Owners, PO, and POs in AMC Digital Library, eResources of the Czech National Library of Technology which cover multiple databases (i.e., SpringerLink, Wiley Online Library, Science Direct, IEEE/IET Electronic library), and Google Scholar. The first evaluation of papers was done by Title and Abstract, and publications with no relation to the PO role were excluded. The remaining publications had the introduction and conclusions reviewed, and those with no valuable information got excluded. The rest of the papers were studied and analyzed in detail. Publications with assumed relation to the PO, which were found cited in papers during the analysis, were extracted and reviewed using similar steps as for initial search. Additional enrichment of the information source and also better connection with practice was achieved by the inclusion of practitioner's books (Cohn, 2010; McGreal & Jocham, 2018; Pichler, 2011; Rubin, 2017; Schwaber, 2004; Schwaber et al., 2015; Schwaber & Beedle, 2002). Due to the lack of papers describing the early adoptions of Scrum, other documents and presentations available for this era were added (Cohn, 2003; Schwaber, 2006, 2009; J. Sutherland, 2007).

OBJ2 - Extract and categorize the activities. Identified papers were read in detail, and activities discovered in text extracted. Those with similar meanings were merged and then divided into categories. There were five categories created: Backlog, Project, Team, Customers and Business, and Development. Activities falling under each category were then chronologically ordered by date of

publication. Although similar activities were mentioned several times in studies, only the first mentioning of the activity was considered for the listing.

OBJ3 - Create the timeline representing the evolution of activities. Identified activities were positioned on the timeline and their change throughout the time discussed. Even here, only the first mentioning of activity was considered to avoid redundancy in the list and to maintain better readability of the timeline.

Validation. Validation of activities identified by the author was performed through the mapping to the activities identified by Unger-Windeler (2019). Activities newly identified during the conduction of this study were listed separately.

3. Product Owner Activities

Jeff Sutherland did the first implementation of Scrum in the software development context at Easel Corp in 1993 (Cohn, 2003). This was the moment when the PO role was created (Sutherland, 2007). In the 1990s, the PO was referred to as the person responsible for managing the Product Backlog to maximize the value of the project. He represented all stakeholders in the project (Schwaber, 2009). Cohn (2003) described that the PO role was usually fulfilled by Product Manager, Marketing, Internal Customer, etc. According to Schwaber (2004), key customers were stepping into the role of Product Owner. Schwaber (2004) was the first who described the role in more detail. Based on Schwaber (2004) main activities of Product Owner can be expressed as: (1) Represent all stakeholders in the project, (2) Maximize the ROI, (3) Cooperate with team sprint by sprint, (4) Define highest priority business value, (5) Create release plans, (6) Create initial requirements, (7) Prioritize the Product Backlog, and (8) Explain value during the team meetings. The importance of the PO role was emphasized by Raithatha (2007), and later more research papers and publications related to PO role followed. Activities from mentioned research papers were used as a source for PO activities extraction and are listed in Figure 1.

To better understand the changes in multiple areas of PO activities, there was a need to categorize them. Therefore, five different categories of activities were defined: Backlog, Project, Team, Customers and Business, and Development. The Backlog category was derived from the definition of backlog management from the Scrum Guide (Schwaber and Sutherland 2017). As the PO was responsible for maximizing the value of the project since the first introduction of the role (Cohn, 2003; Schwaber, 2009), Project was used as one of the categories. The Team was identified as another category because of the mentioned need for cooperation with the development team (Schwaber, 2004). The PO has to represent customers (Schwaber, 2004) and deeply understand the business (Cohn, 2010). These two areas are closely related, so a joined category, Customers and Business, was created. With the PO's direct involvement in the development process (Bass 2015; Bass et al. 2016; Rubin 2017), a new set of activities emerged. Therefore the Development category was introduced.

As the next step, all previously extracted activities were placed into created categories. The results are presented in Figure 1. Activities are chronologically ordered within each category. Each activity record starts with an activity identifier created for further use in this paper and ends with reference to the publication in which the appearance of activity was recognized for the first time.

Figure 1: Categories, activities, and publications; Source: Author, 2019

4. Activities Timeline

All identified activities were positioned on the timeline to provide a view of the change through the lens of time. The PO activities timeline is depicted in Figure 2. The timeline starts with 1995 as this year, the Scrum method, including the PO role, was introduced to the public (Sutherland and Schwaber 1995). Due to the absence of PO activity descriptions in materials before (Cohn, 2003), there were no activities positioned on the timeline before the year 2003.
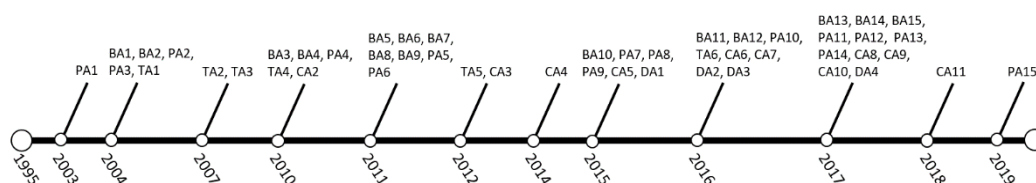
Figure 2: PO activities timeline; Source: Author, 2019

446

For activities in the *Backlog* category (BAn), there was a visible shift identified from setting the initial requirements and giving them a priority (BA1, BA2) to providing a more detailed specification (BA9, BA11, BA14), including acceptance criteria and continuous backlog evolution (BA10, BA12, BA13).

The *Project* activities (PAn) in the timeline indicate that since the beginning, the PO had to maximize the value of the project (PA1, PA2). To do so, the PO needed to start making important decisions (PA4), steer the project (PA5), and manage its economics (PA6), decide about features to be released (PA10) while following plans (PA13, PA14).

The need for close team collaboration (Shammi et al., 2011) was also proved for the PO by activities (TAn) falling under the *Team* category. PO activities in the category evolved from simple cooperation with the team (TA1) through the activities related to the team's motivation and involvement (TA2, TA3), solving conflicts (TA5) to the leadership (TA6).

The PO activities in *Customers and Business* category were originally supposed to be performed by the key customer (Schwaber, 2004), who represented all stakeholders to the development team (CA1). Results show it changed to the activities of an employee who needs to provide the vision (CA2) and manage stakeholder's expectations (CA4). User support (CA8) appeared as a new unique activity.

The results show that the adoption of Agile in large organizations (Dikert et al., 2016; Paasivaara & Lassenius, 2011) brought about new sets of activities, mostly visible in *Customers and Business* and *Project* category. The PO has to gather the requirements (CA5), present the ideas and needs inwards, but also outwards the organization (CA6) and collaborate with stakeholders (CA10). The PO also has to ensure compliance with corporate guidelines (PA7), align more development teams together (PA9), participate in company-wide planning (PA11), focus on process improvement (PA12), and manage the governance (PA15).

Development related activities first appeared in the year 2015, so were considered as the newly emerged category of activities tight to the role evolution with increasing adoptions. The PO got involved in designing, implementation, and dissemination of a reference architecture and providing architecture on large projects (DA1). It was also reported that the PO directly participates in the execution of test activities by testing development outcomes (DA2) or defining test criteria for specified requirements (DA3). The PO also ends the development process by verification of completion of requirements (DA4).

5. Outcomes Validation

The reliability of the findings and correctness of the literature review was validated by comparison with the paper (Unger-Windeler et al., 2019), in which the research team at Leibniz University Hannover did a mapping study on POs in Industry to identify future research directions. As one of the outcomes, they listed PO activities identified during the mapping study. In Figure 3, those activities are mapped to the activities identified in this paper.

The result of the mapping confirmed that nearly all activities from (Unger-Windeler et al., 2019) have corresponding activities identified by the author. The only exceptions were Super secretary and Expert trainer. The reason is that activities came from the papers not related to the PO role, but the On-site customer role from Extreme programming, which was not considered in this study. For Accountability, no equivalent has been identified as the activity definition is very vague. As the rest of the activities or their equivalents were identified in the study, it advocates for not missing any important activities and, therefore, also for the reliability of the study.

Figure 3: Mapping identified activities to (Unger-Windeler et al., 2019); Source: Author, 2020

(Unger-Windeler et al., 2019)	Identified activities
Communication	CA3, PA9
Writing user stories	BA5, BA9
Planning	PA11, PA13, PA14
Prioritize the backlog	BA1, BA6, BA12
Mastering the release	PA3
Technical architect	DA1
Governor	PA7, PA15
Traveler	CA5
Intermediary	CA1, CA3, CA9
Risk Assessor	PA8
Gatekeeper	PA10
Acceptance tester	DA2
Customer relationship manager	CA7, CA11
Managing expectations	CA4
Political advisor	CA3, CA4
Super secretary	-
Visionary	CA2
Accountability	-
Teamwork	TA1, TA4
Expert trainer	-
Critical decision maker	PA4

Moreover, other activities were found that are listed in Figure 1: (BA3), (BA4), (BA7), (BA10), (BA11), (BA13), (BA14), (BA15), (PA1), (PA2), (PA5), (PA6), (PA12), (PA13), (TA2), (TA3), (TA5), (TA6), (CA8), (CA10), (DA3), (DA4). The cause seems to be twofold: (1) Study (Unger-Windeler et al., 2019) did not focus specifically on the PO role activities. In (Unger-Windeler et al., 2019), the main goal was the identification of areas requiring further research, so the search and extraction of activities from different publications were not that thorough. (2) Contrary to this paper, during the conduction of (Unger-Windeler et al., 2019), no books written by consultants from the practice were included. In this literature review books (Cohn, 2010; McGreal & Jocham, 2018; Pichler, 2011; Rubin, 2017; Schwaber, 2004; Schwaber et al., 2015; Schwaber & Beedle, 2002) were included, which provided a new resource for identification of more activities and contributed to the extension of the list of activities.

6. Conclusion

In this paper, the evolution of Product Owner activities since the first introduction of the role was described. As Agile has been spreading through the industry, the PO role has changed from a relatively simple role that was meant to be covered by a customer representative or an internal stakeholder to a more complex role covering a wide range of activities. For the purpose of this study, research papers and publications written by consultants from practice were identified and reviewed. Then, the identified activities were extracted and categorized into five categories: Backlog, Project, Team, Customers and Business, and Development. Within these categories, activities were chronologically listed.

The first activities related to the Backlog were introduced together with the introduction of the PO role and are still PO key activities. Activities in this category shifted from setting the initial requirements and setting priorities to providing detailed specifications while focusing on the continuous backlog evolution. In the Project category, since the beginning, the PO had to maximize the value of the project. To do so, the PO needed to start making important decisions, steer the project, and manage its economics. Alongside the adoptions of Agile in large organizations, the PO has to ensure compliance with corporate guidelines and participate in company-wide planning. Agile software development is done by closely collaborating teams. Team related activities evolved from simple cooperation with the team to the leadership. Activities in Customers and Business category evolved to providing the vision and managing stakeholder's expectations. The PO also has to gather the requirements and present the ideas and needs inwards but also outwards the organization. As a completely new activity, user support has emerged. In development, POs got involved in designing, implementation, and dissemination of a reference architecture, providing architecture on large projects, and participating in test activities. The PO also ends the development process by verification

of completion of requirements. All identified unique activities were depicted in the timeline to represent the outcomes through the lens of time.

This paper provides evidence that the Product Owner role has changed with the agile development expansion. The changes seem to be connected with the increasing adoption of Agile into various types and sizes of organizations, scaling of agile methods, and their tailoring.

Limitations and threats to validity. The research was performed by a single person. It is possible that some of the activities had been missed or incorrectly extracted. To increase the reliability of the paper and to verify the outcomes, identified activities have been mapped to a previous study (Unger-Windeler et al., 2019), which contained the PO activities found during the mapping study conduction as one of the side outcomes. Mapping showed that similar activities for POs except one which was vaguely defined were identified, and in addition, 22 activities have been newly identified. The identification of activities was based only on published resources, and any additional field research has not been conducted.

Further field research will be conducted to validate which of the identified activities of PO are performed in practice and to what extent. There were no factors like geographical locations, size of organizations, or existing frameworks considered in the paper, so findings are valid for a general understanding of the PO role, but not concrete adoptions in specific organizations.

7. Acknowledgment

8. References

Bass, J. M. (2015). How product owner teams scale agile methods to large distributed enterprises. Empirical Software Engineering, 20(6), 1525–1557.

Bass, J. M., Beecham, S., Noll, J., & Razzak, M. A. (2016). All Hands to the Pumps: The Product Owner Role in Small Companies. Lero: The Irish Software Research Centre, 16.

Bass, J. M., & Haxby, A. (2019). Tailoring Product Ownership in Large-Scale Agile Projects: Managing Scale, Distance, and Governance. IEEE Software, 36(2), 58–63.

Buchalcevova, A. (2020). Are Agile and Scaled Agile Frameworks Really Addressing Software Development? IDIMT, 2020, 11.

Buchalcevova, A., & Doležel, M. (2019). IT Systems Delivery in the Digital Age: Agile, Devops and Beyond It. IDIMT, 2019, 9.

Cambridge University Press. (2008). Activity. Cambridge Online Dictionary.

 https://dictionary.cambridge.org/dictionary/english/activity

Cohn, M. (2003). An Introduction to Scrum.

https://www.mountaingoatsoftware.com/uploads/presentations/Introduction-Scrum-SQuAD-2003.pdf

Cohn, M. (2010). Succeeding with Agile: Software development using Scrum. Addison-Wesley, ISBN 978-0-321-66056-5

CollabNet VersionOne. (2019). The 13th State of Agile Survey. https://www.stateofagile.com/

Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations: A systematic literature review. Journal of Systems and Software, 119, 87–108.

Dolezel, M., Buchalcevova, A., & Mencik, M. (2019). The State of Agile Software Development in the Czech Republic: Preliminary Findings Indicate the Dominance of "Abridged" Scrum. CONFENIS, 2019, 12.

Dyba, T., & Dingsoyr, T. (2009). What Do We Know about Agile Software Development? IEEE Software, 26(5), 6–9.

Eloranta, V.-P., Koskimies, K., & Mikkonen, T. (2016). Exploring ScrumBut—An empirical study of Scrum anti-patterns. Information and Software Technology, 74, 194–203.

Kristinsdottir, S., Larusdottir, M., & Cajander, Å. (2016). Responsibilities and Challenges of Product Owners at Spotify—An Exploratory Case Study. In HCSE/HESSD (s. 16).

Lindvall, M., Muthig, D., Dagnino, A., Wallin, C., Stupperich, M., Kiefer, D., May, J., & Kahkonen, T. (2004). Agile software development in large organizations. Computer, 37(12), 26–34.

McGreal, D., & Jocham, R. (2018). The professional product owner: Leveraging Scrum as a competitive advantage. Addison-Wesley, ISBN 978-0-13-468647-9

Oomen, S., Waal, B. D., Albertin, A., & Ravesteyn, P. (2017). How can Scrum be succesful? Competences of the Scrum Product Owner. ECIS, 130–142.

Paasivaara, M., Heikkila, V. T., & Lassenius, C. (2012). Experiences in Scaling the Product Owner Role in Large-Scale Globally Distributed Scrum. ICGSE, 174–178.

Paasivaara, M., & Lassenius, C. (2011). Scaling Scrum in a Large Distributed Project. 2011 International Symposium on Empirical Software Engineering and Measurement, 363–367.

Pichler, R. (2011). Agile product management with Scrum: Creating products that customers love. Addison-Wesley, ISBN 978-0-321-60578-8

Project management institute. (2018). Success in Disruptive Times, Expanding the Value Delivery Landscape to Address the High Cost of Low Performance. PMI. https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pulse-of-the-profession-2018.pdf

Raithatha, D. (2007). Making the Whole Product Agile – A Product Owners Perspective. XP, 184–187.

Rubin, K. S. (2017). Essential Scrum: A practical guide to the most popular agile process. Addison-Wesley, ISBN 978-0-13-704329-3

Shammi, M., Overbeek, S., Verbung, R., Janssen, M., & Tan, Y.-H. (2011). Agile Process for Integrated Service Delivery. Delft University of Technology, 10.

Schwaber, K. (2004). Agile project management with Scrum. Microsoft.

https://pdfs.semanticscholar.org/43e7/ddf062e81253083e2a5b7c619f9bc0d42333.pdf

Schwaber, K. (2006). Scrum et al. Google Tech talk, Google. https://www.youtube.com/watch?v=IyNPeTn8fpo

Schwaber, K. (2009). Terminology in use during the late 1990's. Advanced Development Methods, Inc.

Schwaber, K. (2010). Scrum Guide 2010. Scrum Alliance.

Schwaber, K., & Beedle, M. (2002). Agile software development with Scrum. Prentice-Hall, ISBN 978-0-13-067634-4

Schwaber, K., Hundhausen, R., & Starr, D. (2015). Agile project management with Scrum. Microsoft Press, ISBN 978-0-7356-1993-7

Schwaber, K., & Sutherland, J. (2011). The definitive Guide to Scrum: The Rules of the Game.

Schwaber, K., & Sutherland, J. (2017). The Definitive Guide to Scrum: The Rules of the Game.

https://www.scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf

Sutherland, J. (2007). The Roots of Scrum: How the Japanese experience changesd global software development.

http://jeffsutherland.com/scrum/RootsofScrumJAOO28Sep2005.pdf

Sutherland, J. V., & Schwaber, K. (1995). The SCRUM methodology. OOPSLA Workshop.

Sverrisdottir, H. S., Ingason, H. T., & Jonasson, H. I. (2014). The Role of the Product Owner in Scrum-comparison between Theory and Practices. Procedia - Social and Behavioral Sciences, 119, 257–267.

Unger-Windeler, C., & Klünder, J. (2018). On the Tasks and Characteristics of Product Owners: A Case Study in the Oil and Gas Industry. Product-Focused Software Process Improvement, 3–11.

Unger-Windeler, C., Klünder, J., & Schneider, K. (2019). A Mapping Study on Product Owners in Industry: Identifying Future Research Directions. 135–144.

VALUE ORIENTED BUSINESS MODELING

Małgorzata Pańkowska

University of Economics in Katowice, Poland

Keywords

Abstract

Discussions on objectives of enterprise modeling (EM) come back from time to time in literature and practice. It is particularly visible in promoting new information technology (IT) solution deployment. Usually, IT developers present software functionalities and persuade user that they are useful. However, the questions "why" and "for what" are still without answers. Therefore, information pragmatics for business services processes is main goal of this paper. At first, some IT system evaluation methods are discussed and proposal of business value assessment is included. Later, a case study is to emphasize the information value in approach, which covers BPMN and e3 Value model, and SWOT analysis in ArchiMate language.

1. Introduction

Business information systems (BIS) and enterprise architecture (EA) appear to mainly maintain a practical interest. They include problems that business organizations cope with the current emerging technologies and growing complexity. Particularly EA is understood as a tool for business IT alignment.
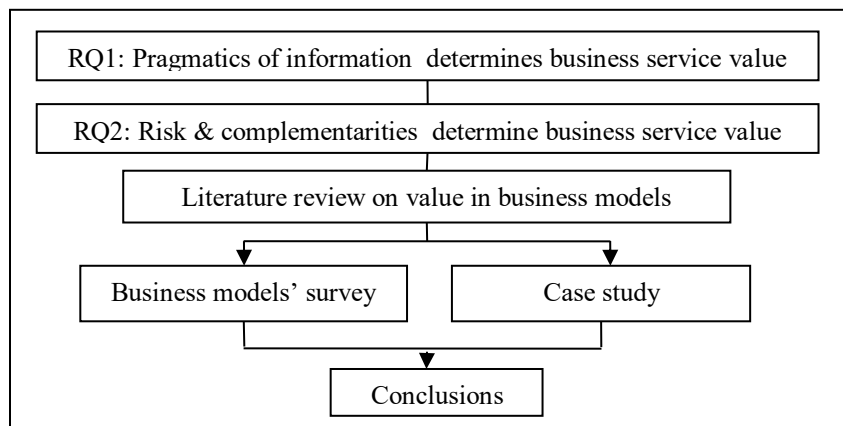
Figure 1. This paper structure

In ISO/IEC/IEEE 42010:2011 standard, EA is understood as consisting of the essential elements of a socio-technical organization and their reciprocal relationships. According to this standard, stakeholder has interest in system that exhibits architecture and has concern that identifies architecture description. This definition does not emphasize the notion of motivation nor value in the EA development process. However, that issues arrive in EA frameworks, i.e., TOGAF or Zachman Framework. Their authors dare undertake discussions on motivation, culture, context, and values in EA development. This paper include literature review and presentation of some evaluation

approaches. Next, the second part include proposal of value analysis and creation method based on risks and opportunities assessments. The last part contains case study for exemplification of the proposed approach. The structure of this paper is presented in Figure 1.

In business science, value is considered as something tangible or intangible, exposed to threat and that is to be secured because of risk of loss. In economics, value is subject of an exchange. Therefore, the RQ1 concerns pragmatics of information, because in the pragmatic dimension, we ask why and for what is the information signifying in business process evaluation. Value of resource is identified with its usability, which means that resource can be accessible and it can be used for established purposes. In this paper, we assume that anything, which loss is painful, has value. Risks depend on values. In RQ2, complementarities are identified with opportunities, business organization strengths and weaknesses. Business value concerns data, information, knowledge, material resources and processes. Value as an abstract is expressed by other characteristics, e.g., compliance, quality, reliability, efficiency, effectiveness, or operability. Many people give priority to economic values, such as prosperity, and emphasize the positive side of technological changes. Other individuals emphasize values such as environmental quality or healthy lifestyle. Beyond value characteristics important for now, there are maintainability characteristics, which determine abilities for future and transferability characteristics, which determine abilities to use or apply a product, service, or information in another different context.

2. Business value models

Values provide orientation for judging and guiding. They are mental constructs allowing individuals and organizations to assess a set of objects and state of affairs (Ronn, 2008). Values are strongly embedded within a cultural context of beliefs, norms and moral convictions. In business, value modeling requires defining models and modeling processes. Process of modeling should be considered as a communication among business stakeholder just to explain what the value is and how it is created. Business concepts included in model are determined by observer of this reality. Each model is presented in a specific modeling language as having a linguistic and a representational function. According to Proper et al. (2019), the linguistic function refers to the ability of a modelling language to frame the discourse about a domain and shaping the observer's concepts in this domain. The representational function refers to the ability of the language to express the conceived domain in a model. This paper survey covered nine models:

- Porter's Value Chain model, which is a collection of activities performed by a company to create value for its customers. Porter emphasizes the issue of integration of roles of customers, employees, and technology for value creation. He discusses the value creation in multi-stakeholder network and in collaborative context (1985). The model is suitable for analysis of traditional manufacturing, but it is less suited for service business (Ward & Daniel, 2006). Unfortunately, the notion of value is limited to the financial dimension, and the activities of the value chain are structured sequentially (Daaboul et al., 2012).

- Business Model Ontology (Osterwalder, 2004), defining the rationale of how an organization creates, delivers, and capture values. In this model, actors are involved in partnerships, which enable providing value proposition. Value configuration determines cost, value proposition – profit, and value channel – revenue (Osterwalder et al., 2005).

- Verna Alle's Value Network Analysis (VNA), including assessment of current and future capabilities for future creation of value (Allee, 2002). This framework considers a value network as a continuously changing systems that reproduces itself. In VNA, value is an

emergent property of the network, so understanding the network functioning is necessary to recognize how and why value is created (Allee, 2008).

- e3 Value Analysis is an approach for modeling value constellations (Pombinho et al., 2014). In relationship networks, actors exchange value objects transferred through value ports, which are directional elements of value interfaces (Von Rosing & Etzel, 2020). A value object is a service, a product or an experience, which is of economic value (Hotie & Gordijn, 2019).

- Resources, Events, Agents (REA) model, explaining actors' exchange of value objects, which are services, products, money, or even consumer experiences. Value modeling includes economic events, in which economic resources change their amounts or features, and the agents, i.e., actors participate in these events (Hunka et al., 2016).

- Value Stream Mapping, which is a lean management tool that helps visualize the steps of product creation process of delivering value to the end-customer (Von Rosing & Etzel, 2020).

- Value Delivery Modeling Language (VDML) models applied to articulate value proposition, activities, and actors (Cummins & de Man, 2011). Value is defined as a measurable benefit delivered to a recipient in association with an item (OMG VDML, 2018).

- Val IT framework is to integrate a set of governance principles, processes, practices and supporting guidelines that help boards, executive management teams and other stakeholders optimize the economic value from IT investments. The concept of value relies on relationships between the expectations of stakeholders and capacities of the resources available to use (Enterprise Value, 2008).

- Enterprise Evolution Contextualization Model (EECM) including the following questions: what, why and how should the enterprise evolve? The EECM components cover concept of enterprise and paradigm of creating value, dimensions, mechanisms and practices (De Vries et al., 2015).

In presented above models, the concept of value relies on the relationships between stakeholders, on transactions of exchanges, on the expectations of stakeholders and capacities of the resources available to use. However, it is necessary to emphasize that value is subjective and determined by stakeholder, i.e., "value is in the eye of the beholder" (Enterprise Value, 2008). Instead of asking about value you can ask about threats, risk and opportunities connected with availability and usage of information, resources, or services. Therefore, it is suitable to ask about the risk of loss.

In a culture of security, an enterprise people are expected to protect information needed to their work. Considering the information as sensitive, usually protected and publicly available help them to establish its value and understand consequences of its use. Through the culture of security people learn value of information. Business resources security management requires business organization context defining, identification of threats, vulnerabilities, risks and opportunities, analysis of all of them, their assessments and treatment, further constant monitoring and planning future proceedings.

In this paper, people are assumed to think about values and valuable resources in aspect of risks. There are risks that must be managed to enhance the opportunity value and the overall value of business or individual activity (Deryck & Huysmans, 2017). As opportunities offer potential gains, it is valuable to interest another party in sharing the efforts to capture a specific opportunity. Notion of opportunity encourages to creativity in resolving business concepts, architectures, designs, and management in social organizations (Forsberg et al., 2004). Opportunities and risks should be managed concurrently, although they require application of separate and unique techniques. In each business organization they are subdivided into and managed as strategic and tactical opportunities and risks. Categories of risk and opportunity are analyzed and managed in a certain business context. Culture and social values within firms as well as in value chain of business partners, or in different

countries constitute the business organization context. Other context characteristics include language, management styles, religious practices, legal, regulatory and reporting requirements, technology support and infrastructure variations.

3. Case study

In this paper, value of information is driven by its use. The same principle is applied for value of services and human competencies. In healthcare services, co-creation of value refers to the interactive nature of services. This co-operation is described as involving a high level of patient participation in the service providing. In co-creation a patient is expected to be active, and provides input to service provider before, during and after the service. Patient as well as other stakeholders, e.g., physician, nurse, technician can be considered as value creators. In general, healthcare services aim providing reliable services for patients, but there are other goals, i.e., continual self-improvement, tackling huge challenges and risk, balancing responsibility among patients, employees, and society, integrity and ethics in all aspects of practice. Beyond the service goals, there are also business purposes, i.e., increase visibility of service costs, tracking and documenting the healthcare processes, more accurate diagnoses and therapies, facilitation of drug evaluation process, reduction of administrative costs through reliable IT implementation. Each objective is to be interpreted in terms of the benefits that it could produce for different healthcare stakeholders.

The purpose of this case study is to facilitate understanding the value creation in healthcare services by summarizing an extension of the process-based view. We define and present in Business Process Modeling Notation (BPMN) a business process of patient admission to hospital (Figure 2). Further, we develop a conceptual model of co-creating value in healthcare service network (Figure 3), that integrates different stakeholders, i.e., patient, physicians and nurse. BPMN and e3 Value models present an idea that value is co-created by multiple parties involving a symbiotic relationship between a hospital and its primary stakeholders, i.e., patients. The value cannot be created in isolation. An information, material and human resource (i.e., human competency) have value if they enable a hospital unit to realize performance and achieve expected goals. Rare resources and competencies provide more advantages. Other characteristics of advantageous objects are as follows: appropriatability, inimitability, non-substitutability, and imperfect mobility. Human competencies, i.e., knowledge, engagement, and professional skills are factors that can lead to service inimitability.

Next, through the application of value-focused process reengineering to risk management models, we develop model that extends the capabilities of existing e3 Value and BPMN models and enables risk-oriented value management. The proposed model (Figure 4) is illustrated in the context of healthcare services in a general practice hospital to demonstrate the practical application of and the resulting benefits from the ArchiMate language use.

The ArchiMate value streams (Figure 4) can be decomposed into other value streams, which can be used to represent value stream stages. Each of the value stages can be associated to value. Beyond that, in ArchiMate language the value streams can be related to the business processes and outcomes, which represent the value proposition. In Figure 4, values are realized through the use of business process architecture. Values are for the stakeholders, particularly for patients. Business process improvement by constant renewing of strengths-weaknesses-opportunities-threats (SWOT) assessment is the key issue needed by the organization, based on a clear understanding of the business mission and vision. Assuming that healthcare services are realized for and with patients, the values, i.e., reliability, effectiveness, validity, safety, security and efficiency, are patient-oriented. In the aspect of hospital management staff, value items could be different and oriented towards economic efficiency.
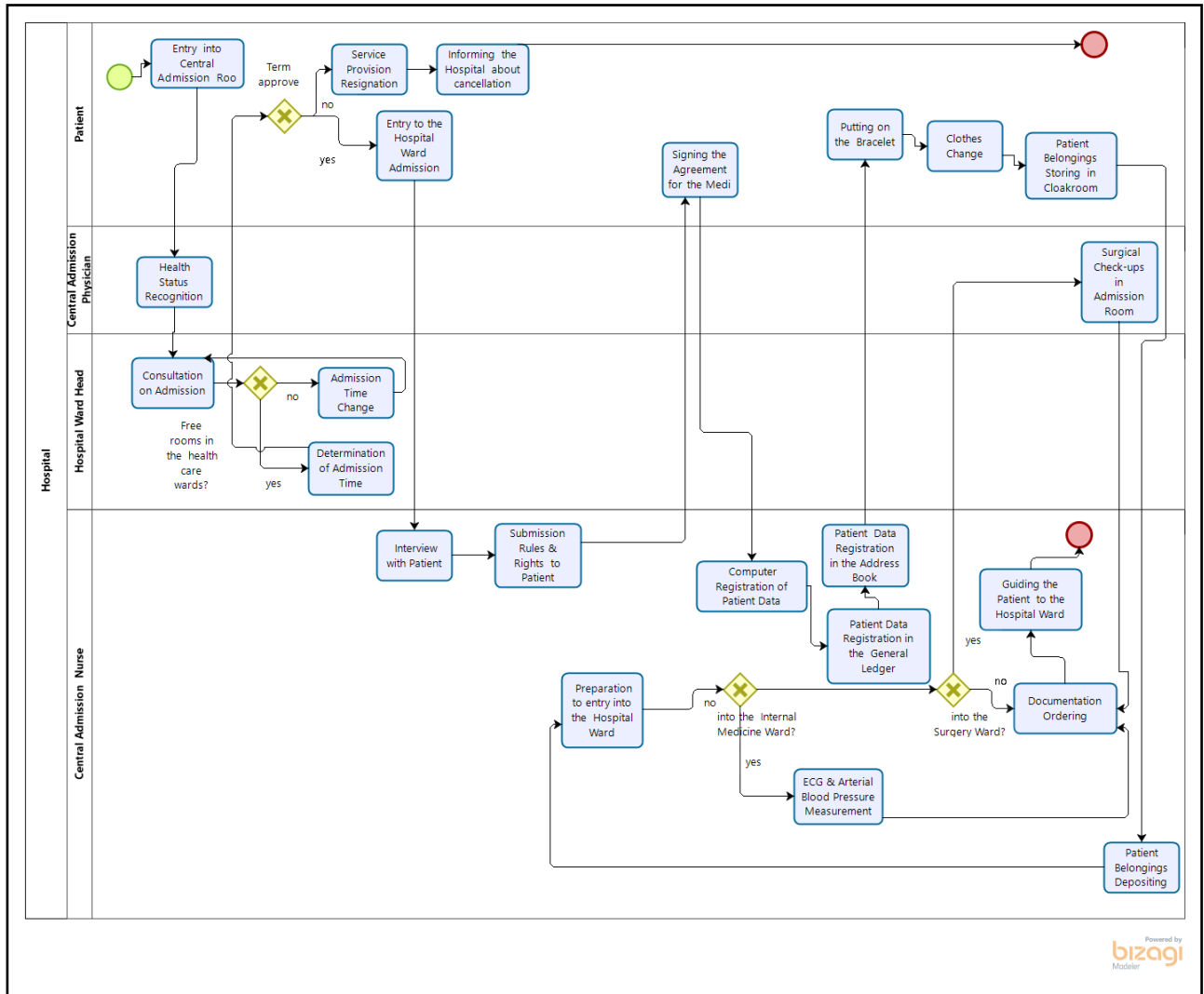
Figure 2. Patient hospital admission process in BPMN

In this paper, we integrate the hospital ecosystem conceptualization with existing modeling approaches. Our results how limited support for value creation analysis. The e3 Value model presents network of relationships between stakeholders, but basing on this network analysis further concluding on values is required.

The BPMN business process modeling focuses on stakeholders' tasks, cycles, documents and their mutual dependencies, therefore concluding on values is not directly supported. However, researchers are still able to develop techniques that can provide a useful, conceptual, and tooling basis for value creation analysis.
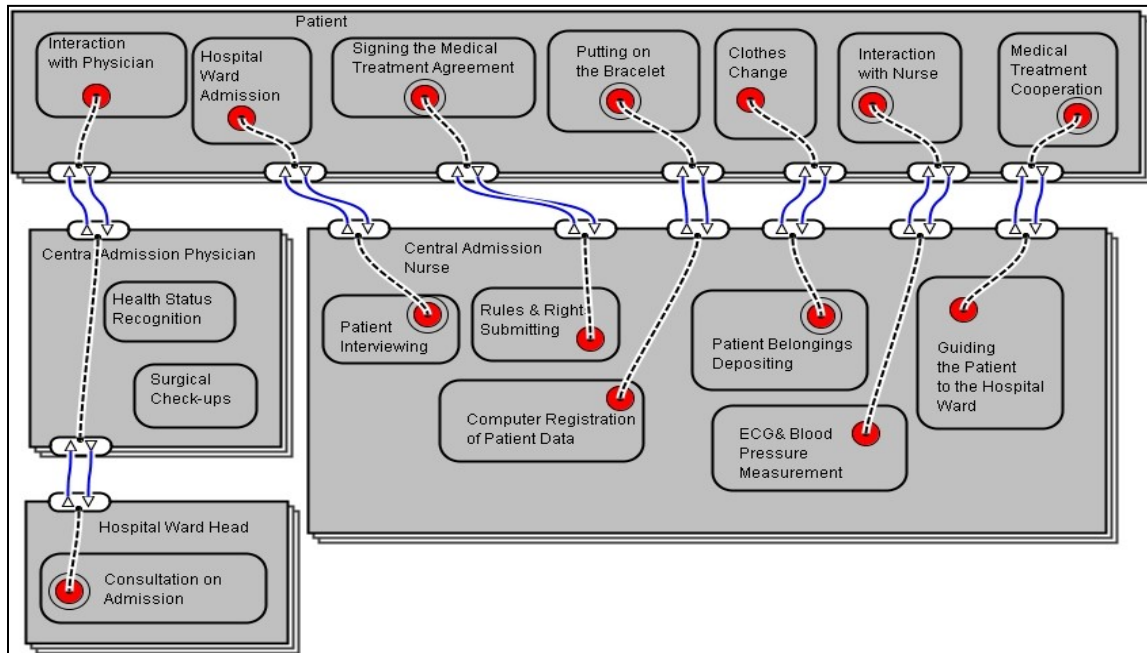
Figure 3. Patient hospital admission value network in e3 Value notation
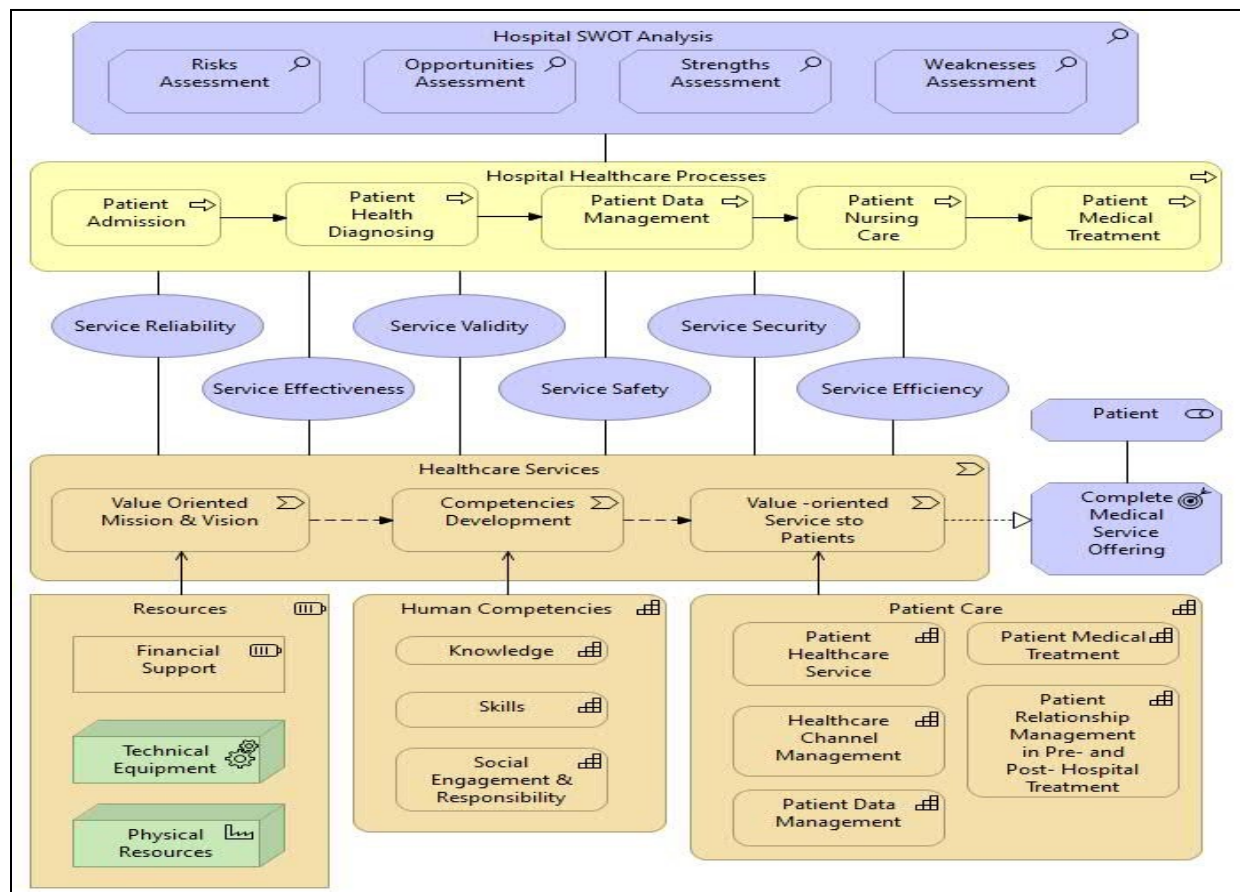
Figure 4. Healthcare service values model in ArchiMate

Just ArchiMate language and ArchiMate value modeling technique are expected to enable the understanding, communication, and analysis of value creation. As it was presented above, values rely on stakeholders' interests' alignment and requires the understanding of ecosystem risks and complementarities, i.e., opportunities, organizational strengths and weaknesses. Values are not so

clearly identifiable, therefore, the better solution is to leave them intangible, but certainly determine conditions and context, in which they are realized.

4. Conclusions

The task of value recognition is crucial for business organizations. Information pragmatics, as an integral part of information semiotics concerns the intention of relationships between information and the human behavior in a specific social context. These relations and communication are presented in BPMN and e3 Value models, which are developed to approve that information pragmatics determine healthcare service value. The ArchiMate model describes how the task of value creation can be supported by impact of risks and opportunities. In Figure 4, values are presented as a distinctive operational goals. Value creation in healthcare ecosystems relies on a specific structure of stakeholders' mutual alignment. The ArchiMate model is to support the research question that risks and complementarities determine business service value. In general, value modeling techniques can improve understanding of how ecosystem determines value creation and the alignment of structures and processes development. High performing in risk and opportunity recognition can permit to conclude on value creation processes.

5.   References

Allee, V. (2002). The Future of Knowledge: Increasing Prosperity through Value Networks. Butterworth- Heinemann, Boston, ISBN 0-7506-75918.

Allee, V. (2008). Value Network Analysis and value conversion of tangible and intangible assets. Journal of Intellectual Capital Volume, 9, No 1, 5-24.

ArchiMate Enterprise Architecture Modeling Language (2020) The Open Group. Accessed April 2020 at https://www.opengroup.org/archimate-forum/archimate-overview

BPMN, Business Process Model and Notation (2011). Object Management Group. Accessed April 2020 at https://www.omg.org/spec/BPMN/2.0/

Cummins, F.A., de Man, H. (2011). Capability Analysis with the Value Delivery Modeling Language, Capitalizing on capabilities, Cutter IT Journal, vol 24, no 4 , 33-39. Accessed April 2020 at http://www.cutter.com .

Daaboul, J., Castagna, P., Bernard, A. (2012). From Value Chains to Value Network, Modeling and Simulation, 9th International Conference on Modeling, Optimization and Simulation. Accessed April 2020 at https://hal.archives-ouvertes.fr/hal-00728567 .

De Vries, M., Gerber, A., van der Merwe, A. (2015). The Enterprise Engineering Domain, Conference: Enterprise Engineering Working Conference, June 2015. Accessed April 2020 at https://www.researchgate.net/publication/283554111_The_Enterprise_Engineering_Domain/figures?lo=1

Deryck, M., Huysmans, P. (2017). Application of Enterprise Engineering to Lean Process Management: An Explorative Case Study. In Aveiro, D., Pergl, R., Guizzardi, G., Almeida, J.P., Magalhaes, R., Lekkerkerk, H. (eds.) Advances in Enterprise Engineering XI, 141-148, Springer, Cham, ISBN 978-3-319-57954-2.

e3 Value, The e3 value methodology: Exploring innovative e-business ideas (2020). Accessed April 2020 at https://research.e3value.com/e3family/e3value/

Enterprise Value Governance of IT Investments, The Val IT Framework 2.0. (2008) IT Governance Institute, Rolling Meadows. Accessed April 2020 at http://www.isaca.org .

Forsberg, K., Mooz, H., Cotterman, H. (2004). Vizualizing project management, models and frameworks for mastering complex systems, Wiley &Sons, Inc. Hoboken New Jersey. ISBN-13: 978-0471648482.

Gordijn, J., Osterwalder, A. Pigneur, Y. (2005). Comparing Two Business Model Ontologies for Designing e-Business Models and Value Constellations, BLED 2005 Proceedings. Accessed April 2020 at http://aisel.aisnet.org/bled2005.

Hotie, F., Gordijn, J. (2019). Value-Based Process Model Design. Bus Inf Syst Eng, 61(2), 163-180.

Hunka, F., van Kervel, S.J.H., Matula, J. (2016). Towards Co-creation and co-production in Production chains modeled in DEMO in REA Support. In Aveiro D., Pergl R., Gouveia D (eds.), Advances in Enterprise Engineering X, 54-68, Springer, Heidelberg, ISBN 978-3-319-39567-8.

ISO/IEC/IEEE 42010:2011. System and software engineering – Architecture description (2011). Accessed May 2020 at https://www.iso.org/standard/50508.html

OMG VDML, Value Delivery Modeling Language Version 1.1. (2018). Object Management Group. Accessed April 2020 at https://www.omg.org/spec/VDML/

Osterwalder, A. (2004). The Business Model Ontology. A proposition in a Design Science Approach, Universite de Lausanne, Ecole de Hautes Etudes Commerciales. Accessed April 2020 at http://www.hec.unil.ch/aosterwa/PhD/Osterwalder_PhD_BM_Ontology.pdf.

Pombinho, J., Mendes, C., Fragoso, B., Santos, R., Silva, N., Sixpence, E., Tibolet, J. (2016). Objectifying Value Co-creation - An Exploratory Study. In Aveiro, D., Pergl, R., Gouveia, D. (eds.) Advances in Enterprise Engineering X, 37-53, Springer Heidelberg, ISBN 978-3-319-39567-8.

Proper, H.A., Bjekovic, M., van Gils, B., Hoppenbrouwers, S.J.B.A. (2019). Towards a Multi-stage Strategy to Teach Enterprise Modelling. In Aveiro, D., Guizzardi G., Guerreiro, S., Guedria, W. (eds), Advances in Enterprise Engineering XII, 181-202, Springer, Cham, ISBN 978-3-030-06097-8.

Renn, O. (2008). Risk Governance, Coping with Uncertainty in a Complex World, Earthscan, London, ISBN-13: 978-1-844-07291-0.

The TOGAF Standard, Version 9.2 Overview (2018). The Open Group. Accessed April 2020 at https://www.vanharen.net/the-togaf-standard-version-9-2/

Ward, J., Daniel, E. (2006). Benefits Management Delivering Value from IS &IT Investments, John Wiley & Sons, Chichester, ISBN-13 978-0-470-09463-1

Von Rosing, M., Etzel, G. (2020). Why a value-oriented process modelling concept is needed, 41-48. Accessed April 2020 at http://CEUR-WS.org/Vol-2574/short4.pdf.

Zachman, J.A. (2008). The Concise Definition of The Zachman Framework. Zachman International Enterprise Architecture. Accesses April 2020 at https://www.zachman.com/about-the-zachman-framework.

CRITICAL SUCCESS FACTORS IN BIG DATA PROJECTS

Bára Smolová

Faculty of Economics
Technical University of Liberec
bara.smolova@tul.cz

Keywords

Big Data, Critical Success Factors, CFSs, project management, Key factors, data volumes, successful project.

Abstract

The amount of data grows exponentially every year. Organisations are aware of the wealth contained in this data and continue to make greater investments into Big Data technologies as a result. Such investments are intensive in terms of the time required and financing involved, and often do not achieve the expected results. Many Big Data projects result in failure when the project itself is not finished or is completed with significant deficiencies. The root cause of such a high level of failure remains unclear. A large number of critical success factors have been defined for Big Data projects, but these have failed to increase the success rate substantially. The purpose of this paper is to identify those key critical success factors that currently exist and that may be considered critical and not just important. The research is carried out through a questionnaire survey and semi-structured interviews with respondents who have many years of Big Data experience. A total of five basic critical success factors for Big Data projects were identified.

1. Introduction

Big Data are considered to be large repositories of variously structured data that can be acquired, aggregated, stored and analysed in order to obtain new information and knowledge (Manyika et al., 2020). The volume of such data increases exponentially every year (Reinsel et al., 2018). This growth is driven by new technologies (sensors, smart devices, etc.) and increases in the volume of content on the Web (social networks, etc.).

The need to process data to generate the necessary insights and information from such data increases as the overall total volume of data increases. The solution is to use Big Data technologies, which offer a variety of options for using available data. Interest in these solutions grew the most in 2011 and has remained high ever since then (Google Trends, 2020).

Despite the tremendous interest devoted to this issue, projects with such focus often end in failure. According to an official statement from Gartner, 60% of all Big Data projects are unsuccessful (Gartner, 2016). According to a later, unofficial statement from one of Gartner's analysts, D. Heudeker, the actual rate is 85% (Heudeker, 2017). Other sources state that 55% of Big Data projects are incomplete and many others fail to achieve the defined requirements (Kaskade, 2013). Such a high level of failure is alarming but the root cause has yet to be determined.

A great deal of research focused on identifying critical success factors has been completed to improve the success rate of Big Data projects. A large number of diverse factors, in which it is easy to get lost, were the outcome of these studies. The objective here is to select those factors that may be considered truly critical, meaning those factors with the greatest influence on the success of Big Data projects.

2. Theoretical framework

The first mention of Big Data dates to 1997, when NASA employees characterised large quantities of data as a tremendous challenge for computing systems (Cox and Ellsworth, 1997; Friedman, 2012). The systems of the day were unable to visualise the quantities of data that were needed. Such a quantity of data was called Big Data (Cox and Ellsworth, 1997).

Currently, Big Data may be characterised through the basic three "Vs" (3Vs) as described by Douglas Laney, a Gartner data analyst, in research conducted in 2001 and this definition also added capital letters to the properties of such data: Volume (large quantities of data), Velocity (the ability to process as quickly as it is generated) and Variety (data in various structures) (Laney, 2001).

Veracity (data from reliable sources) was added to these characteristics of Big Data in 2011 resulting in the 4Vs (Snow, 2012). Value (valuable data for business) followed suit in 2012 (IBM, 2012).

Other characteristics have accumulated over time. The most commonly used definition of Big Data is the mentioned 3Vs: Volume, Velocity and Variety (Oracle, 2020).

Technologies focused on processing Big Data are indispensable in many aspects of human life.

2.1. Critical Success Factors (CSFs)

With the goal of increasing the success of Big Data projects, various research has been conducted to identify the critical success factors of Big Data projects (Yeoh & Popovič, 2016; Evers, 2014; Gao et al., 2015; Koronios et al., 2014; Saltz & Shamshurin, 2016; Wamba et al., 2020; McAfee & Brynjolfsson, 2012; Gómez & Heeks, 2016; Chen et al., 2016, Alharthi et al., 2017).

Overlapping factors were aggregated into ten basic factors for the purposes of providing an elementary assessment of critical success factors in this research:

1. **Definition of goals and requirements** (business cases, strategies, vision, missions and project documentation)

    (Yeoh & Popovič, 2016; Evers, 2014)

2. **Qualified project team** (experience, expertise, abilities and personal traits of team members)

    (Gao et al., 2015; Yeoh & Popovič, 2016; Evers, 2014, Koronios et al., 2014, Saltz & Shamshurin, 2016)

3. **Project management** (management methods and the like)

    (Gao et al., 2015; Koronios et al., 2014; Saltz & Shamshurin, 2016)

4. **Change management** (suitable processes, etc.)

    (Yeoh & Popovič, 2016; Saltz & Shamshurin, 2016, (Wamba et al., 2020)

5. **Data quality and availability** (source systems)

    (Gao et al., 2015 5, Yeoh & Popovič, 2016, Evers, 2014, Koronios et al., 2014, Saltz & Shamshurin, 2016; McAfee & Brynjolfsson, 2012; Gómez & Heeks, 2016; (Wamba et al., 2020)

6. **Corporate culture** (hierarchies, approaches, such as data-driven, barriers, level of communication across the company)

(McAfee & Brynjolfsson, 2012, Alharthi et al., 2017)

7. **Solution development model** (selection of a model, such as cascade, iterative, etc.)

(Yeoh & Popovič, 2016)

8. **IT infrastructure** (the scalability, security and flexibility of systems)

(Yeoh & Popovič, 2016; Evers, 2014; Saltz & Shamshurin, 2016, Gómez & Heeks, 2016;

9. **End user participation in project implementation** (reaction to changes during the project)

(Evers, 2014)

10. **Suitable technologies** (selection and usage)

(Chen et al., 2016; Gao et al., 2015; Evers, 2014; Koronios et al., 2014; McAfee & Brynjolfsson, 2012; Wamba et al., 2020)

The five critical success factors with the greatest impact on the success of Big Data projects were identified from these in the questionnaire survey.

3. Data collection and methodology

The defined objective of this research was to identify those factors that had the greatest impact on the success of Big Data projects and therefore those which are truly critical.

The research used a combined quantitative and qualitative methodology. Quantitative research involved the use of a questionnaire survey and responses were received from 42 respondents with experience involving Big Data projects.

Qualitative research was conducted through an in-depth semi-structured interview. For this portion of the research, only those respondents with long-term experience involving IT projects with a special focus on Big Data were contacted. The first part of the interview was structured, while the second part was not structured so as to obtain more detailed information and pertinent opinions on the topic at hand.

42 respondents participated in a questionnaire survey and another 3 semi-structured interviews.

3.1. Data collection

The questionnaire survey was conducted electronically and through in-person and telephone-based querying. Those respondents involved in the semi-structured interview were queried in person and online (using Skype or other means).

A total of 45 respondents participated in the survey, and the number of potential respondents itself is significantly limited given that Big Data is a specific niche in the IT industry and only a few people are specifically focused on it.

4. Results

The results of both research approaches are presented in this section.

4.1. Quantitative research

A questionnaire survey involving 42 respondents was conducted from 2019 to 2020. The obtained data is depicted in Figure 2 and Figure 3.

The questionnaire survey contained 25 questions and was focused on the experience respondents had with Big Data projects, and specifically the success of such projects, critical success factors, team size, the applied technologies and more.

Figure 1 shows success rate of Big Data projects according to the questionnaire survey

The pie chart on the left shows that the failure rate is around 69%.

The pie chart on the right shows a closer breakdown of the responses of respondents who claim that their last Big Data project did not end successfully. 47.6% of respondents state that they did not meet all the set requirements, while 21.4% that they did not meet most of the set requirements of the project.
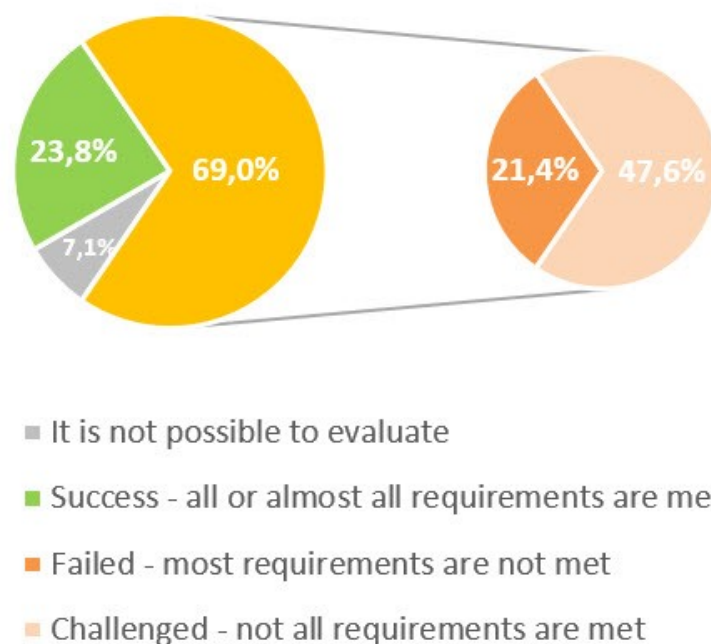
Figure 2: Success rate of Big Data projects according to the questionnaire survey

The respondents were asked to select the five most important of the critical success factors defined in Chapter 2.1. The survey results are presented in Figure 2.

Figure 2 shows which of the factors defined above are considered by the respondents to be the most critical - a qualified team, availability and quality of data, appropriate technology, definition of project objectives and requirements, as well as project management. The development model is considered to be the least important factor.

Figure 3: Success of Big Data projects according to evaluation of Critical Success Factors

The survey also focused on other aspects of Big Data projects but they are not included in this paper.

4.2. Qualitative research

A total of three respondents participated in this portion of the research and completed the semi-structured interview related to their experience with Big Data projects.

The structured portion of the interview included 10 questions focused on Big Data-related activities of the respondents themselves.

From the interview answers were selected to the following question: In your opinion, what are the critical success factors for a Big Data project?

"I see input data and the transparency of that data as the primary critical success factor of a project. I would rank an experienced team second, and the ability of team members to understand data in context and to know how to work together is critical, while the experience and skills of team members are also important. I often encounter situations where requirements are unclear at the very beginning of a project or where a project budget and timeline are established without any connection to the

quality of the available data. These things need to be clarified before starting the project" Ing. Jana Hamalová, Senior Analyst

"I consider the specifications of customer requirements and expectations, data quality, the experienced project team of the solution supplier and, of course, communication across the company's departments to be critical factors in the success of Big Data projects. It is also very important what tool is used in the project. The partnership of the supplier with the vendor company is also very welcome" Bc. Karel Tvrzník, Splunk administrator.

"The success of a Big Data project is the preparation of clear, reachable and valuable use cases before deployment because of the significant investment of hardware or cloud resources, and top management expectations" Yasmin Lak, Big Data Managed Services team lead.

Respondents were not informed in advance about the examined success factors, they were only asked what, in their opinion and experience, are the critical factors for the success of Big Data projects.

Respondents who were subjected to a semi-structured interview largely agree with the results of quantitative research.

5. Conclusions and future work

Quantitative research showed that respondents ranked a qualified project team, the availability and quality of data, the use of suitable technologies, the definition of objectives and requirements and project management as the most important critical success factors in the success of a Big Data project. The questionnaire survey also verified that the failure rate of these projects is very high.

The respondents in the qualitative research defined similar critical success factors. The key factors for the success of a project include a qualified team made up of experienced people with the necessary skills and abilities to work with data and who are able to cooperate with one another. This was followed by the use of suitable technologies that meet all the requirements of the given project. Finally, precise, clear and timely specification of the project was identified as important, including use cases and customer requirements. The defined requirements should be achievable within a given time and financial framework.

The main limitation was to find respondents who have experience with Big Data projects and are willing to fill out a questionnaire or undergo a semi-structured interview. The situation was exacerbated by the ongoing COVID-19 pandemic, when the willingness to answer the questionnaire was even lower.

This research focused on the issue of critical success factors in Big Data projects will continue and the form of future research will be adapted based on the results obtained within this work with the goal of obtaining more detailed data for more in-depth analysis.

6. Acknowledgement

7. References

Cox, M., & Ellsworth, D. Application-Controlled Demand Paging for Out-of-Core Visualization. In NAS. NAS. https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf.

Evers, J. M. (2014). Critical Success Factors Of Business Intelligence And Big Data Analysis: Utilizing hidden Business value from Big Data for Business Intelligence [Master Thesis]. Tilburg University.

Gao, J., Selle, S., & Koronios, A. (2015). Towards A Process View on Critical Success Factors in Big Data Analytics Projects. In Twenty-first Americas Conference on Information Systems: AMCIS 2015 (pp. 1-14). Association for Information Systems.

Gartner. (2017). Through 2017, 60% of #BigData projects will fail to go beyond piloting. In Twitter. Twitter. https://twitter.com/gartner_inc/status/718423161289732098?lang=cs (Accessed: 12 April 2020).

Gómez, L. F., & Heeks, R. (2016). Measuring the Barriers to Big Data for Development: Design-Reality Gap Analysis in Colombia's Public Sector. In Development Informatics: Working Paper Series (62nd ed., pp. 1-13). University of Manchester.

Google Trends. (2020). Big Data. Google Trends. Retrieved May 02, 2020, from https://trends.google.com/trends/explore?date=2010-01-01%202020-01-05&q=big%20data (Accessed: 1 May).

Heudeker, N. (2017). We were too conservative. The failure rate is closer to 85%. And the problem isn't technology. In Twitter. Gartner. https://twitter.com/nheudecker/status/928720268662530048?ref_src=twsrc%5Etfw (Accessed: 1 May).

Hoeren, T., & Döpke, C.Kolany-Raiser, B. (Ed.). (2018). Big Data in Context: Legal, Social and Technological Insights (January 2018). Springer. https://link.springer.com/book/10.1007%2F978-3-319-62461-7.

Chen, H. M., Schutz, R., Kazman, R., & Matthes, F. (2016). Amazon in the Air: Innovating with Big Data at Lufthansa. 2016 49Th Hawaii International Conference On System Sciences (Hicss), 5096-5105. https://doi.org/10.1109/HICSS.2016.631.

IBM. (2012). The Four V's of Big Data. Retrieved February 13, 2018, from http://www.ibmbigdatahub.com/infographic/four-vs-big-data (Accessed: 1 May 2020).

Kaskade, J. (2020). Survey: What IT Teams Want Their CIOs to Know About Enterprise Big Data. In PR Newswire: press release, distribution, monitoring. PR Newswire Association. https://www.prnewswire.com/news-releases/survey-what-it-teams-want-their-cios-to-know-about-enterprise-big-data-188190311.html (Accessed: 14 February 2020).

Koronios, A., Selle, S., & Gao, J. (2014). Big data project success - A meta analysis. In Pacific Asia Conference on Information Systems (pp. 171-209). PACIS.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. In Gartnet Blog Network. Gartner. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

Manyika, J., Chui, M., Brown, B., & Bughin, J. (2020). Big data: the next frontier for innovation, competition, and productivity. Digital Insights, 1. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation (Accessed: 18 January 2020).

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. Harvard Business Review, 2012, 60-69.

Oracle. (2020). What Is Big Data? The Definition of Big Data. Oracle: Integrated Cloud Application And Platform Services. Retrieved May 01, 2020, from https://www.oracle.com/big-data/what-is-big-data.html (Accessed: 25 April 2019).

Reinsel, D., Ganz, J., & Rydning, J. (2017). Data Age 2025: The Evolution of Data to Life-Critical, Don't Focus on Big Data; Focus on the Data That's Big. In Seagate: Storing the world's digital content. IDC White Paper. https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf.

Reinsel, D., Gantz, J., & Rydning, J. (2018). Data Age 2025: The Digitization of the World From Edge to Core.: An IDC White Paper. Data Age 2025, 1-28. https://www.seagate.com/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf.

Saltz, J. S., & Shamshurin, I. (2016). Big Data Team Process Methodologies: A Literature Review and the Identification of Key Factors for a Project's Success. In International Conference on Big Data (pp. 1-8). IEEE. https://doi.org/10.1109/BigData.2016.7840936.

Snow, D. (2012). Adding a 4th V to BIG Data - Veracity. In Dwaines Snows Thoughts on Databases and Data Management. Blogger. http://dsnowondb2.blogspot.com/2012/07/adding-4th-v-to-big-data-veracity.html. (Accessed: 25 April 2020).

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2020). How'big data'can make big impact: Findings from a systematic reviewand a longitudinal case study. International Journal Of Production Economics, 2015(165), 234-246. https://doi.org/10.1016/j.ijpe.2014.12.031.

Yeoh, W., & Popovič, A. (2016). Extending the understanding of critical success factors for implementing business intelligence systems. Journal Of The Association For Information Science And Technology, 67(1), 134-147. https://doi.org/10.1002/asi.23366.

HUMAN AND CULTURAL IMPACTS ON PROJECT SUCCESS

HUMAN AND CULTURAL IMPACTS ON PROJECTS

Gerhard Chroust

Johannes Kepler University Linz
gerhard.chroust@jku.at

Keywords

Abstract

Successful international cooperation and globalization in project teams is based on effective communication. National and cultural differences (such as language, habits, taboos, norms, and world view) cause obstacles and misunderstandings to surface during communication. In this paper we describe several current models for describing the cultural dimensions of nations and their different (apparent and hidden) impacts on behavior, beliefs, attitude and project work. We discuss insights and approaches for successfully handling resulting obstacles.

1. Cultural Differences

Globalization and international cooperation have brought more intense and diversified contacts between persons of different cultures. Travelling to foreign countries and meeting people of other cultures we often experience surprise, or even a 'cultural shock' (fig. 1), while observing their behavior and hearing their arguments. This begins with simple things like counting one's fingers, basic conventions as to perform business transactions, plus ultimately deeply rooted beliefs, e.g. about women's role in society. We understand 'culture' as the shared complex system of language, value system, norms, religion, myths, beliefs, manners, behavior, and structure which is characteristic of a society or part of it.

Fig. 1: Apparent national differences

2. Layers of Communication and Cooperation

Each type of communication between humans usually involves several layers, some obvious, some hidden, some easy to change, some deeply imbedded in the human nature. Fig. 2 shows a 7-layers model of human communication, where higher layers have increasing cultural dependency and sensitivity (Chroust, 2008).
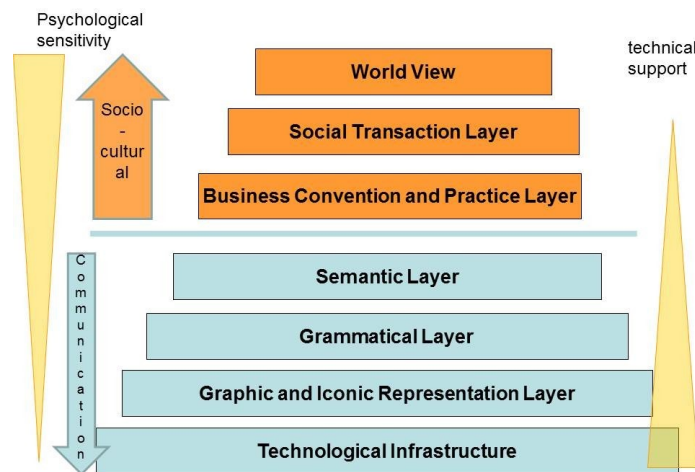
Fig. 2: Layers of communication/cooperation

Technological Infrastructure Layer: It provides the technical basis for communication (telephone, e-mail, internet, ...). Most of it is invisible to the users, but provides the equipment for connecting people: representation of characters (including national characters and two-byte characters (Adams, 1993)), reading direction (left-to-right, right to left) (He et al., 2002; Kim, 1999), provision of sufficient storage space for data, sorting algorithms, appliance of the locale which defines the proper representation of date, currency, time, etc. (He et al., 2002; Kubota, nd).

Graphic and Iconic Representation Layer: Increasingly software products rely on graphical representations in panels, demos, and animations. The correct symbolic meanings of signs, colors, animals, etc. has to be taken care of. When presenting persons on a screen the body language(!) and local and private setting is very critical. Human viewers are highly sensitive to mismatch and discrepancy.

Grammatical Layer: Text created by a computer program (be it an original or a translation) must be correct including jargon, local habits, and idiosyncrasies. Even within one language there are sometimes discrepancies resulting from different social standing, educations, etc.

Semantic Layer: The meaning of words and sentences taking into consideration ambiguity, high-context vs. low-context culture (Meyer, 2016), etc. have to be dealt with. The difference between technical and common language, expressiveness of languages, ambiguities, abbreviations and jargon also needs attention.

Business Conventions and Practices Layer: This layer comprises issues such as leadership approach, organizational structure, jurisdiction, deadlines and commitments, overtime, acceptable performance measures, etc. Discrepancies may cause misunderstanding, mistrust, alienation, and loss of business opportunities.

Social and Communication Layer: What are the forms and styles of communication with respect to politeness, replying to a customer or user, expressing criticism (Meyer, 2016), using gender-adequate language, ways to express 'no', metaphors puns, jargon, and humor.

World View Layer: This layer summarizes deep-seated beliefs and feelings dealing with aim and purpose of one's life, religion, social classes, social positioning, position of women, racism, taboos, etc.

The last three layers are intrinsically related to differences in cultural points of view. In order to undertand their implications basic dimensions of cultural preferences in different nations must be understood.

3. Modelling Cultural Differences

From 1970 onward computer networks and electronic mail have gradually provided easy communication between other nations and have also started to influence international business and commerce. As a consequence cultural differences have also been recognized as causes for misunderstandings and have for this reason also reeived growing scientific interest. Serious attempts to describe and understand cultural differences have started soon, encouraged by international trade and cooperation. G. Hofstede coined the term 'software of the mind' (Hofstede & Hofstede, 2005). Since then many different model were designed and offered as a means to describe and explain cultural differences and analyze the culture-related behaviors of people. Three different approaches to modelling methods can be identified (fig. 3) .
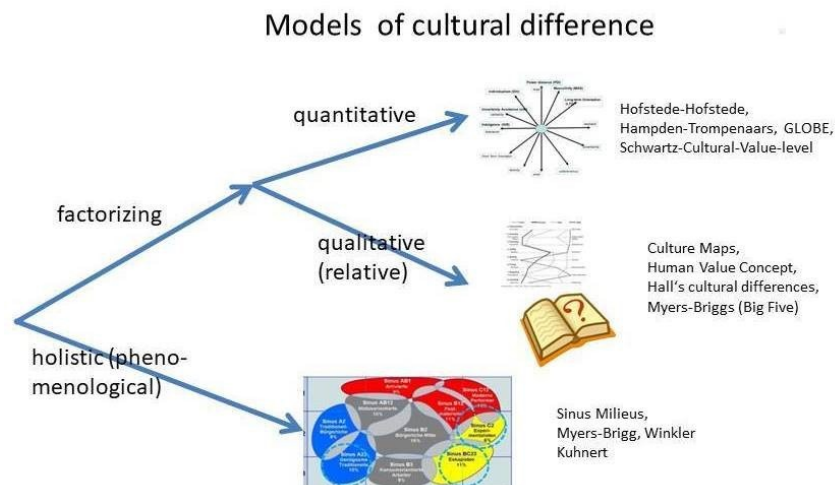
Fig. 3: Differing descriptive Models

The factorizing approach identifies several (usually 4 to 9) differing dimensions of basic cultural orientation and assigns different values ('scores') to the different cultures. The totality of the values of these dimensions characterizes a certain (large !) segment of the total world population. The chosen dimensions are expected to be more or less orthogonal, i.e. that the value in one dimension does not imply the value for another dimension. Only (Hofstede et al., 2010) explicitly tries to prove orthogonality of the dimensions in his model by showing that for each pair of dimensions the values are not correlated. We distinguish two sub methods:

Quantitative models assign numerical values ('scores') to each dimension allowing numerical comparison (e.g. section 4.1, section 4.2, and section 4.3.)

Qualitative (relative) models specify only a relative ordering between individual nations with respect to their suspected scores section 4.4.

The holistic (phenomenological) approach - nased on a judgement of the basic behavior a small number of distinct groups are idnetified by considering their lifestyle and behavior . (see section 4.5).

3.1. Warnings and Caveats

When using cultural mappings the limits of such models have to be considered.

- Humans take part in different social groups and often play different roles. Depending on the role people might show different behavior : the brutal boss might be a most loving father at home.

- No matter into how many different segments a world will be divided, there will always be a huge number of divergent personalities in each segment. For example, if we speak of the behavior of 'Han-Chinese', we speak of a group of approx. 1,3 billion persons. There must be considerable differences within this group.

- The data give just an average and should not be understood as a stereotype for a group of people. The scores are expected to obey a distribution curve most likely similar to a Gaussian distribution (see fig. 4).

Fig. 4: Average versus stereotype

4. Well-known Models

Numerous models have been proposed in the last 50 year, only a handful has been accepted and widely used. We will present 4 of the best known ones.

4.1. Hofstede-Hofstede

Having started with 4 dimension G. Hofstede presented 2010 a 6-dimensional models having (Hofstede et al., 2010), see fig. 5. Except for 'Indulgence/restraint' they all found their way into the GLOBE-model (section 4.3).

Power Difference Index (PDI): This is the extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally. The inequality is endorsed by the followers as much as by the leaders.

Individualism Index (IDV): This is an indication of how loose the ties between individuals are as compared to (collectivistic) societies in which individuals from the moment of birth are integrated into strong, cohesive in-groups, often extended families which continue to protect them in exchange for unquestioning loyalty.

Masculinity (MAS): This characterizes how big the gap in the behaviour of men and women is by only by following the values linked to their gender: i.e. men's values (assertiveness and competitiveness) versus woman's' values (modesty, care, consolation, etc.).

Uncertainty Avoidance Index (UAI): This deals with a society's tolerance of uncertainty and ambiguity, of how uncomfortable the individual feels in unstructured situations which are novel, unknown, surprising, or different from the accustomed. To avoid uncertainty strict laws, rules, safety and security regulations are enforced. On the philosophical and religious level certainty is gained by a belief in an Absolute Truth.

Long-Time Orientation Index (LTO): LTO stands for the fostering virtues oriented toward future rewards by thrift and perseverance in contrast to Short Term Orientation fostering of virtues related

to the past and the present - in particular with respect to tradition, fulfilling of social obligations, the keeping of countenance, and "saving one's face"'.

Indulgence versus Restraint (IVR): Indulgence stands for a society that allows relatively free gratification of basic and natural human behaviour related to enjoying life and having fun. Restraint stands for a society that suppresses gratification of needs and regulates it by means of strict social norms.
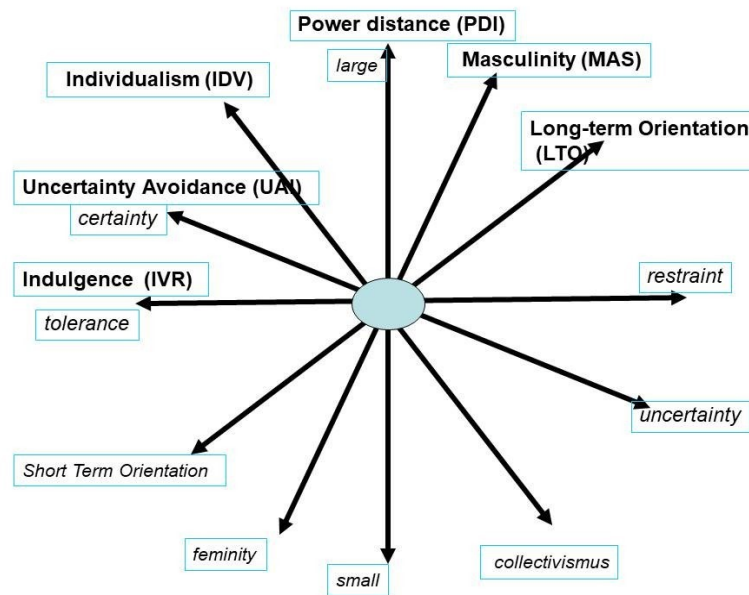

Fig. 5: Hofstede's 6 cultural dimensions

4.2. Hampden-Trompenaars

Ch. Hampden-Turner and F. Trompenaars (Hampden-Turner & Trompenaars, 2000) based their model on data similar to G. Hofstede /citeHofstede-10 but deduced 7 widely different dimensions.

universalism-particularism Universalism is about finding broad and general rules. Particularism is about judging the case on its own merits,

individualism-communitarianism Individualism is about the rights of the individual, while Communitarianism is about the rights of the group or society.

specificity-diffusion Focusing on the specific role of a person or situation or object versus looking at them holistically.

achieved status - ascribed status Achieved status is about gaining status through performance while ascribed status is about gaining status through other means, such as seniority or birth.

inner direction-outer direction Inner-directedness concerns thinking and personal judgement 'in our heads'. Outer-directed is seeking data and information in the outer world.

sequential time-synchronous time Time as sequence sees events as separate items in time, sequenced one after another. Time as synchronization sees events in parallel, synchronized together.

4.3. The GLOBE Model

The GLOBE Model ('Global Leadership and Organizational Behavior Effectiveness', 1994-2014, (House et al., 2004)) was lead by Robert J. House. He at first focused on leadership, but soon

branched out into other aspects of national and organizational cultures. This model includes essentially the models from G. Hofstede and to some extent from Hampden-Trompenaars.

Tab 1: 9 dimensions of the Globe Model (House et al., 2004)

Globe Dimension	from (see note)	
Uncertainty avoidance	HH, HT	
Power distance	HH	
Institutional collectivism	HH HT	from 'individualism'
In-group collectivism	HH HT	from 'individualism'
Gender egalitarianism	HH	from 'masculinity'
Assertiveness	HH	from 'masculinity'
Future orientation	HH	from 'longterm orientation'
Performance orientation	none	the degree to which innovation, high standards, and excellent performance are encouraged and rewarded.
Humane Orientation	none	the degree to which fairness, altruism, generosity, and kindness are encouraged and valued
Indulgence vs. Restraint	HH	not used
Universalism vs.	HT	not used
Achieved Status vs.	HT	not used
Inner direction vs. outer	HT	not used
Sequential Time	HT	not used
	Note. H: (Hofstede et al., 2010), HT: (Hampden-Turner &Trompenaars, 2000)

4.4. Meyer - Cultural Map

Erin Meyer (Meyer, 2016) chose a different approach: She identified 8 dichotomic pairs of essential communication behaviour:

Tab 2: Meyer's Culture Map

communication behavior	range of behaviour
Communicating	low context vs. high context
Evaluating	direct negative vs. indirect negative feedback
Persuading	principles-first vs applications-first
Leading	egalitarian vs. hierarchical
Deciding	~~consensual vs. top down~~
Trusting	based on task vs. relationship
Disagreeing	confrontational vs. avoiding confrontation
Scheduling	structured vs. flexible

4.5. Sinus Milieus

Sinus Milieus (Barth & Flaig, 2013) start with observing and a ppreciating the total ('holistic') world view of a person. The model identifies 11 different milieus, differentiated as shown in fig. 8,

resulting in a chart like fig. 9.

- social situation
    - o lower middle class / lower class
    - o middle middle class
    - o upper class / upper middle class
- basic orientation
    - o tradition : (hold on) (preserve).

o modern individualization (having & enjoying), (being & changing).

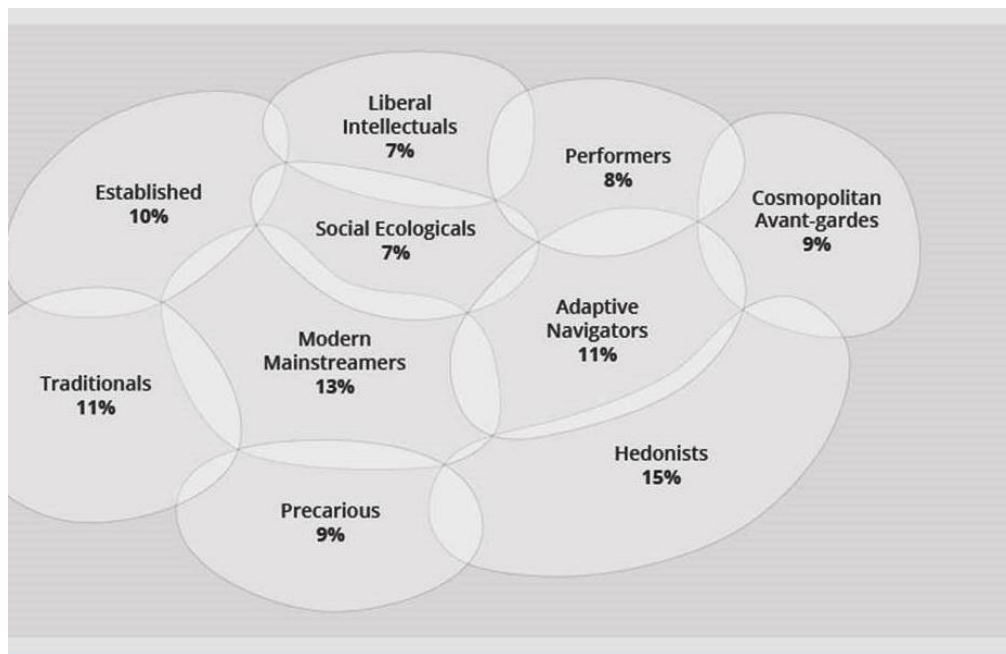o reorientationization : (doing & experiencing) (overcoming borders).

Fig. 6: Typical chart of Sinus Milieus

5. Consequences for Project Cooperation: Cultural Proficiency

Due to globalization the cooperation between different nations is becming more and more intense. Cooperation with persons from other cultures create difficulties and misunderstanding which often create frustration and anger and may be detrimental to the project success. When two persons communicated they expect a certain (culturally dependent!) code of behavior based on a certain amount of a common world view.

What is said is always interpreted in the cultural context of the adressed person. Most of this is unconscious, especially if both partners belong to the same cultural environment. The higher up in the hierarchy of fig 2 the more apparent culture differences are felt and the more difficult it is to bridge them.

The explicit content and form of the message is only part of the total communication. Body language plays a major role in face-to-face situations (Decker, 1992; Molcho, 2006; Morris, 1994) but also in man-machine communication. A certain level of Cultural Proficiency can avoid situations of this kind and is a prerequisite in the cooperation with people from other cultures.

Tab 3 shows 6 levels of cultural proficiency, its impact on human behaviour and on the design of software-intensive systems.

Tab 3: Cultural Proficiency and Software-intensive Systems

level	human behaviour	software-intensive-system design
Cultural destructiven-ess	see the difference, stomp it out, eliminate other people's cultures.	force users to follow the prescribed communication and concepts even if contrary to their cultural expectation or precondition
Cultural incapacity	see the difference, handle it wrongly, belief in the superiority of one's own culture and behaviour that disempowers another's culture.	'correct' or 'improve' on culturally relevant interactions

Cultural blindness	see the difference, act like you don't, act as if cultural differences you see do not matter, or not recognizing that there are differences among and between cultures.	believe that you 'know' how to include cultural variation, but just use traditional cliches and ignore any discrepancy
Cultural pre-competence	see the difference, respond inadequately. Awareness of the limitations of one's skills or an organization's practices when interacting with other cultural groups.	provide certain cultural parameters which the user is encouraged to change, but the support it inadequately
Cultural competence	see the difference, understand the difference that difference makes.	all relevant parameters of cultural difference are adjustable and flexible, the responses are adequate
Cultural proficiency	see the differences and respond positively and affirmingly. esteem culture, learn about individual and organizational culture, and interact effectively.	It is doubtful whether a system would/could reach this level. It would mean interactively recognizing cultural aspects to adapt and learn.

(Lindsey et al., 2003)

6. Outsourcing

Outsourcing is a business practice in which a company hires a third-party to perform tasks, handle operations or provide services for the company. IT outsourcing is a booming business. In cross-border cooperation it makes good economic sense for both parties, since it promise reduced cost, improved performance and access to a wider labor market. Due to cultural difference, however, it is not an unproblematic process.

Outsourcing of software production, especially in Asian countries, does need international cooperation and communication across national/cultural borders. Problems can arise both when designing a product and also when performing a development process (Kobayashi-Hillary, 2005) (Krishna et al., 2004). Of even more importance is the cultural proficiency (see fig. 10) on both sides of the outsourcing project. Successful outsourcing needs a careful selection of the type of product to be produced - imbedded software is less culturally dependent than user interface design.

7. Localization

Global sales via internet efforts prove problematic as they show people in their 'natural environment', however, this is highly culturally dependent on lifestyle, habits, and expectations. The growing sophistication of computer interaction (thanks to Artificial Intelligence) is of the effect, that software interfaces are enriched by more and more human-like features (Barbour & Yeo, 1996; Chroust, 2000). The users intuitively attribute certain human properties to these interfaces- and expect the software to behave consistently according to their cultural expectation (their cultural background!). Or as (Chroust, 2007) states: 'it should behave like a courteous, understanding butler'.

We speak of 'localization' :"the process of adapting a product to reflect the local standards, culture and language of another market. "

If the (culturally dependent!) body language does not convey the right message, the spoken word will not be believed and the message not taken notice of: "you have to be believed to be heard" (Decker, 1992).

Important decisions are

- Choice of languages as target

- Potenial candidates for translation Planning documents for the project, Installation Instructions, Operating/Handling Procedures, Instructions for adaptation and minor modification, Maintenance Instructions (?), Customizing Instructions, End-user Manual

- Which set of all candidate documents really have to be translated and which one is not. Typically operator instructions of an English original might not need translation, since the operators are expected to have sufficient command of English.

- Packaging and Charges: Delivering one source product with all available language option or delivering individual language products. This includes decisions on the charging strategy.

7.1. Language Translation

An important part of localization is the translation of the involved languages, but as the previous section argues, not the only one!

Several methods exist for Language Translation.

7.1.1. Human translators

They have existed since ancient time. It is advisable for translators to have the target language as their mother tongue, in order to be able to adjust the translation to the cultural requirements of the target language and to avoid smaller and larger pitfalls, which might create problems or embarrassment. This approach is costly, lengthy and depends on the varying human abilities.

7.1.2. Computer based translation

Computer-based translation is the (largely) automatic conversion of texts from one language into another language. While human translation is part of applied linguistics, machine translation is a sub-area of Artificial Intelligence. Several methods exist

Rule-based translation This largely mimics a human translator: The translation program generates a syntax tree (based on the syntax of the source language) for the source text. It converts the words in the source into words in the target language, using traditional dictionaries, but also tries to resolve amgibuities etc. From the source syntax tree a target syntax tree is derived with the translated words. Obvious further improvements are made to the target text.

Example-based machine translation The example-based translation relies on a rich corpus of bi-lingual phrases. Each input phrase is compared to existing exemplary source language phrases. The most similar one is used as base for the translation. The quality of the translation depends on the amount and the quality lf the available phrases for translation, based on huge amounts of learning material [Kirkpatrick-20].

The advantage of this method is that companies having to update their product documentation, manuals, guidelines, etc. periodically can carry losts of text over into the next version.

Statistical machine translation Statistical machine translation is based on statistical models, the parameters of which are derived from the analysis of bilingual text corpora. The statistical translation systems must be based on large amounts of available texts in reference translations. whciha re used to train the translation system (Wikipedia-english, 2013, keyword: 'machine translation').

Neural network machine translations Neural networks are trained with huge amounts of bilingual texts. The systems are self-learning, mostly without human help. The translation process is noclonger transparent.

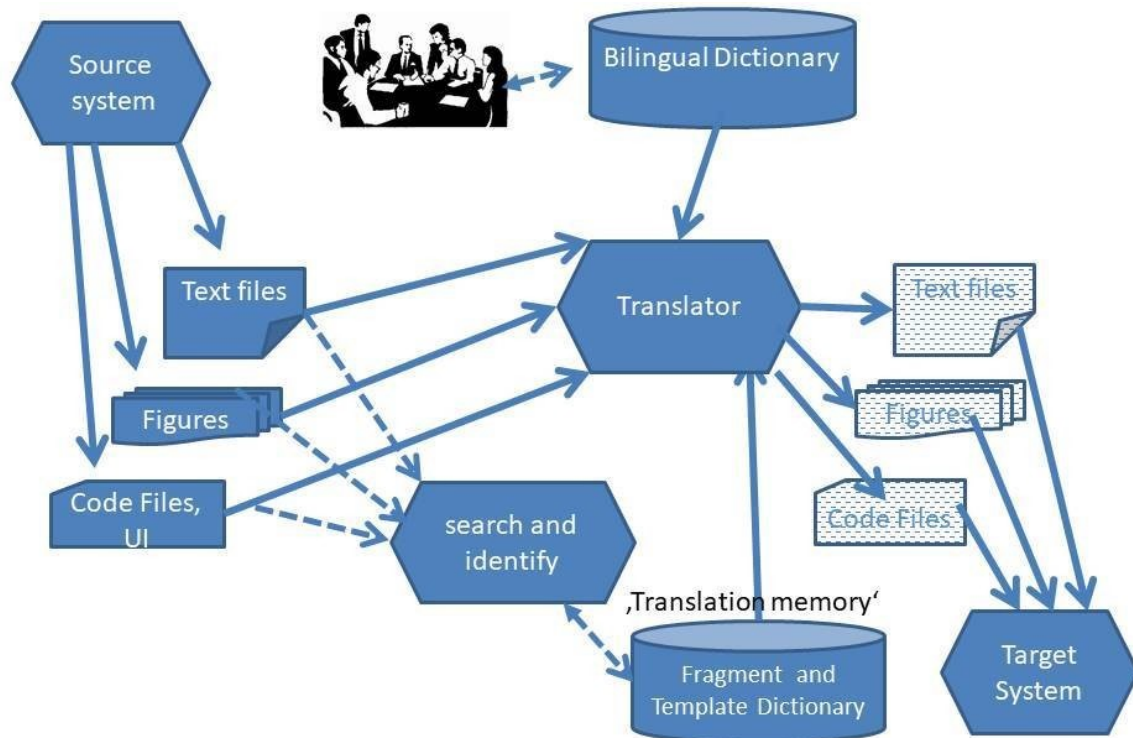A typical over-all translation process is shown in Fig. 7.

Fig. 7: The overall translation process

7.1.3. Commercial 'back box'

Many commercial software houses offer translation services in the form of an 'app, a 'black box' (e.g. Google Translate). The applied methods are usually a corporate secret, but they can be expected to rely on all of the methods above. For small amounts of text and occasional users this is very quick, easy-to-use, often a cheap or cost-free option.

8. Summary

Studying cultural difference is a highly interesting subject of growing importance, triggered by the globalization of business, lifestyles and travel. In the business world cultural differences might make or break a company's success.

9. References

Adams, G. (1993). Internationalization and character set standards. StandardView, 1(1):31–39.

Barbour, R. H. & Yeo, A. (1996). Bilingual/multi-lingual business software: the spreadsheet- in Proc. 1996 Information Systems Conf. of New Zealand (ISCNZ'96), pages 63–71.

Barth, B. & Flaig, B. (2013). Was sind sinus-milieus? - eine einführung in die sozialwissenschaftliche fundierung und praxisrelevanz eines gesellschaftsmodells. In Thomas, P. M. & Calmbach, M., editors, Jugendliche Lebenswelten. Berlin - Heidelberg 2013.

Chroust, G. (2000). Internationalization is more than language translation! In Hofer, S., Beneder, M. (eds.): IDIMT-2000, 8th Interdisciplinary Information Management Talks, Verlag Trauner, Linz, pages 431–440.

Chroust, G. (2007). Software like a courteous butler - issues of localization under cultural diversity.

In ISSS 2007. The 51th Annual meeting and Conference for the System Sciences. Tokyo, Japan, August 5-10, 2007. Curran Associates, Inc. (Mar 2008).

Chroust, G. (2008). Localization, culture and global communication. In Putnik, G. D. & Cunha, M. M., editors, Encyclopedia of Networked and Virtual Organizations, vol II, pages 829–837. Information Science reference, IGI Global, Hershey USA 2008.

Decker, B. (1992). You've got to be Believed to be Heard. St. Martin Press, New York 1992

Hampden-Turner, C. & Trompenaars, F. (2000). Building Cross-Cultural Competence - How to Create Wealth from Conflicting Values. Yale Univ. Press 2000.

He, Z., Bustard, D. W., & Liu, X. (2002). Software internationalisation and localisation: practice and evolution. In PPPJ '02: Inaugural conference on the Principles and Practice of programming, 2002, pages 89–94, Maynooth, County Kildare, Ireland. National University of Ireland.

Hofstede, G. & Hofstede, G. J. (2005). Cultures and Organizations - Software of the Mind. McGraw-Hill, NY 2005.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). Cultures and Organizations - Software of the Mind, 3rd edition. McGraw-Hill, NY 2010.

House, R., Hanges, P., , Javidan, M., P.W., D., & Gupta, V. (2004). Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies. Sage Publ. Calif., 2004.

Kim, K. (1999). Standardizing romanization of Korean Hangeujl and Hanmal. Computer Standards and Interfaces, vol. 21 (1999), no. 5, pages 441–459.

Kobayashi-Hillary, M. (2005). Outsourcing to India. The Offshore Advantage. Springer Berlin 2005. Krishna, S., Sahay, S., & Walsham, G. (2004). Managing cross-cultural issues in global software outsourcing. Comm. ACM, 47(4):62–66.

Kubota, T. (n.d.). Introduction to i18n. http://www.debian.org/doc/manuals/intro-i18n/, [Feb. 2003]. Lindsey, R. B., Robins, K. N., & Terrell, R. D. (2003). Cultural Proficiency: A Manual for Schoo Leaders. Corwin Press 2003.

Meyer, E. (2016). The Culture Map : Decoding How People Think, Lead, and Get Things Done Across Cultures. Public-Affairs 2016, Kindle Edition.

Molcho, S. (2006). ABC der Körpersprache. H. Hugendubel Verlag, Kreuzlingen/München 2006. Morris, D. (1994). The Human Animal. BBC Books, London 1994.

Wikipedia-english (2013). Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/.

GENERATION Z AND PREFERENCE OF MESSAGES WITH SHORT-TERM IMPACT

Zdeněk Vondra

Department of multimedia, Faculty of informatics and statistics
University of Economics, Prague
zdenek.vondra@vse.cz

Keywords

Abstract

This paper refers to generation Z's focus and preference on short-term impact messages in communication stream consumption. This effect was observed on online communication channels content connected with high school students' decision making about their future study on university. Findings in this paper are based on qualitative research done by in depth interviews and survey research. This research brought discovery that high school students prefer gathering information about their short-term future perspective. Short-term focused communication content was applied in a development and execution of communication campaign that brought significant results of nearly 43% raise of applicants on undergraduate study of information and data science. After elimination of other influences, it proves the concept. Results of research and its output application were discussed in perspective on changing behavior of youths caused by social media and their content's ultimate accessibility which causes addiction on short messages and their ability of dopamine reward. Conclusion of these findings is the need for careful acceptance of this state implementation of short-term related principles in information and communication systems development and especially in their interfaces.

1. Introduction

Aim of this paper is to refer to the outcomes of research and application of its results both focused on the problem of effective setting of communication content for youths. Benefit of this finding should be found about communication or interface making for youth (generation Z) recipients or users. This paper refers to experience based on communication with youths of generation Z focused on arousing their interest in the study of quantitative oriented undergraduate programs. Our team observed the effect of messages following their short-term goals. The messages focused on describing the near future and explaining fulfilling their needs in a short-term perspective have a bigger impact on their decisions about future than long term-oriented ones. This situation happens although their decision will majorly affect their life career direction. The effect was observed in a stream of online communication primarily using social media. Social media are described as media that allows its users to make and publish communication content (Obar & Wildman, 2015). Communication that brought the explained experience was based on social interaction and brought the social media features to normally linear communication channels. The observed

effect is discussed from the perspective of young people's communication habits connected with dopamine circuitry addiction caused by social networks.

This paper reflects findings discovered in the research from 2018 focused on finding out the reasons why high school students are not interested in occupation in insurance industry or insurance mathematics in Czech Republic (Hanzalová et al, 2018). Findings from this research were used as the base for further research on high school students that focused on applicants to undergraduate degrees in information and data science. Findings from both researches were applied in a communication campaign in 2019 with the intention to bring more applicants on undergraduate degrees in information and data science. Communication campaigns changed the standard of communication content focused on faculty benefits, success stories and expertise. Content was changed to short-term issues like living in dormitories, party occasions and study gadgets in the perspective of the first year of study. Communication campaign was executed by online marketing channels with high importance of social networks especially Facebook and Instagram. The results of the communication campaign were a significant raise of applicants to nearly 43% more than the year before.

2. Generation Z

Understanding the upcoming generations in the meaning of productive age and society direction major impact is important in the way of intergenerational communication, sharing values and opinions and motivation acquiring. Actual generations of employers solve many situations associated with motivation and work habit issues connected with generation Y and Z. Generation Z demographically means people born in the late 1990s and early 2000s. One of the most important characteristics of this generation is that they are used to using digital media and social networks as a natural medium of their communication and a source of information (Seemiller & Grace, 2016).

The MML-TGI analysis highlights the importance of describing the generation of millennials that precedes Generation Z, not only in terms of age, but especially in terms of lifestyle, attitudes, preferences, or values that manifest across generations. Millennials are described as strong Internet users, users of social networks, often own smartphones, educated, interested in the environment, reject traditional gender roles, are hedonistic, travel in their free time, like to get to know new cultures, change jobs more often, prefer teamwork, the main source of information for them is the Internet, TV plays a small role, they are open, they like news and changes, they buy online more, they take care of the origin of products (fair trade), they organize holidays themselves, they do not use travel agencies, they do not travel repeatedly to the same places (MEDIAN, 2017). According to the Digital News Report 2017 (Newman, 2017), respondents in the Czech Republic most often use social media for communication not only for access to news, but for any other purposes. The most important medium for respondents in the age group 18-24 is clearly Facebook (82%). In second place was YouTube, which is used by 71% of respondents. The third most used medium is Facebook Messenger (64%). Other social networks are already used significantly less. Instagram is used by 38% of respondents, followed by WhatsApp (22%), Snapchat (19%), Viber (15%), LinkedIn and Google Plus (5%), Line and WeChat (2%). The following generation Z is corresponding with the last years of secondary schools thus falls within this definition.

As mentioned professor Ahmed in his paper focused on smartphone and social media usage (2019) more than half of generation Z spend daily nine hours or more using smartphone and regard to the use of social media by generation Z nearly one-third of the students admitted using social media for 7-10 hours per day, and slightly less than one-third of the students spent 5-6 hours per day on social media.

3. Research methods

Findings in this paper are based on qualitative research done in 2018 by in depth interviews and survey research. Aim of the qualitative research was originally to uncover the students' interest in insurance industry and insurance mathematics in Czech Republic (Hanzalová et al, 2018). Research was focused on awareness about occupation, decision process, motivations, and communication channels. During the research there was found that high school students have no awareness about insurance occupations. So the research was turned forward uncovering the relation to study information and data science study programs because it was the closest study field that students can identify as their choice and can probably direct them to study insurance or insurance mathematics in the future. The qualitative research was therefore suitable for the purpose of this paper. The in-depth interviews were done with 25 respondents from the terminal grade of high schools all over the Czech Republic to represent the overall situation. Selected respondents had an assumption to study information and data science oriented undergraduate study programs. The goal of the research was to understand main motivations, the decision processand process and a major way of obtaining information about their university programs selection. Interviews were done in two - membered teams in scope from 45 to 60 minutes. Every interview was deeply reviewed.

Quantitative part of the research was based on secondary research consisting of analysis of surveys Eurostudent VI (Fischer & Vltavská, 2016), MML-TGI: analysis of millenials (MEDIAN, 2017), Digital news report 2017 (Newman, 2017), Demographic evolution and projection of university performance (Koucký & Bartušek, 2011) and High school graduates and the labor market: ECONOMIC INDUSTRIES (Doležalová & Vojtěch, 2016).

Findings from in depth interviews and secondary research were followed by questionnaire surveys with 175 high school student respondents. This survey was also focused on awareness about future occupation, decision process of selection undergraduate study, motivations, and communication channels. The importance of the research was to confirm and link the findings from the previous qualitative and secondary research.

4. Results of the research

Eurostudent research found out that only one third of students were long term decided what study program they will apply on undergraduate study and they can adapt to it (Fischer, Vltavská, 2016). Undecided two thirds are making decisions at the last minute. Respondents realize that the choice of further study is a great substantial decision. However, respondents choose university studies they deal with later, and much less intensively and responsibly than they originally intended. This results in reluctance and low motivation to really address the topic (even among respondents who seemed to be responsible). Shows respondents want to satisfy their need to find a suitable university degree with a minimum effort. Selection is preceded by a decision process based on passive gaining a portfolio of options and the consequent narrowing of possibilities.

Decision process of undergraduate study selection revealed by in-depth interviews consists of three phases:

1. passive building of knowledge,
2. passive consideration of options,
3. active consideration of options.

Building of knowledge means unintended passive building of awareness about suitable undergraduate study options or just directions of possible career. Passive consideration means unintended passive evaluation of study options which makes attributes and personal ranking of selected study opportunities. Active consideration means getting rid of options from previous phases. Knowledge builds the whole high school (1st - 3rd year), during this period they passively think about possible variants (nothing pushes them at that moment and so they add alternatives). In the graduation year, they make active decisions, but they are under strong pressure from teachers, parents, classmates, universities, and the media. In addition, they solve the school-leaving examination and the university is another big task for them in addition to the school-leaving examination. Final selection is based on criteria of amusement, impression of good investment, future application possibility, earnings, practicality, reasonable difficulty of study, location. These findings reveal the process as waiting for information about study possibilities without active discovery. Active role of a student is the reviewing of the offered options which puts him in the role of decision advisor, but it is his own very fundamental decision, which affects his future life. This indicates that the student prefers comfort and things that come to fulfilling his own will.

Subsequent questionnaire surveys that's importance was to confirm or specify the findings from secondary research and in-depth interviews brought the key insight that for the current high school graduate, the shape of a new life at university is at least as or more important than the field he / she will study. This brought new light to the whole process and confirmed the passive role and importance of comfort. Current candidates from the generation Z, in situations where it is a crucial decision of their lives, prefer short-term comfort and avoid thinking about long-term personal goals.

Key knowledge about social networks mentioned in the part focused on generation Z description was confirmed. It makes sense to communicate via Messenger, Facebook, and YouTube. Instagram rather does not have to pay off. It is a channel where most current and potential students are available. University communication on social networks should be lightened, it should be consistent with the fact that the target group wants to have fun.

5. Application of research findings

University team used the revealed insight that for the current high school graduate, the shape of a new life in college is at least as or more important than the field he / she will study for a setup of a new communication campaign of faculty. The campaign brought a fundamental change in communication of undergraduate study programs to high school students in terminal grade. Therefore there was not spoken about study programs or faculty but about the life of a university student in the first semester with the key message focused on comfort, life change and shape of new life at university via micro influencers on social media. The message was "You know where you're going with us, we're a community of people like you, just a year or two away, so contact us and we'll help you handle it."

Result of application rise was that the campaign achieved an increase of 42.8% in comparison with previous year and the goal of the campaign was thus more than quadrupled (original expectations was 10% rise). The achieved result not only stopped the long-term decline in applications, but also a significant jump, which is noticeable at the level of the entire university, which increased the number of applications thanks to the faculty result. For elimination of other possible effects causing this result, there was studied the change in application count on other faculties of the university and collected the feedback. Change in application count on other faculties of the university was from 4% decrease to 5% increase. The feedback from the applicants and students was also on behalf of right set communication content.

6. Discussion

These researches brought the discovery that high school students prefer gathering information about their comfort which is connected to their short-term future perspective. Short-term focused communication content distributed via social media and micro influencers was applied in a development and execution of communication campaign that brought significant results of nearly 43% raise of applicants on undergraduate study of information and data science. After elimination of other influences, it proves the concept that short-term oriented communication content can be very effective also in affecting the decision process about significant life steps of generation Z.

The main reason for this state is the effect of usage of social media by generation Z that is from 5 to 10 hours a day at two thirds of the population as mentioned in the previous parts of the paper (Ahmed, 2019). This topic is spoken a lot in public all over the world with revealing of an addiction behavior. In 2015 there was also developed the Social Media Addiction Scale (SMAS) which quantifies the level of addiction (Tutgun-Ünal & Deniz, 2015). Social media addiction leads to change in behavior which is explained by the dopamine loop. Dopamine is delivered to the brain based on stimuli which provides the feeling of quick rewards that is very addictive. As we get used to quick rewards, we are impatient for more difficult rewards to obtain and our mind starts to focus on short-term actions with rewards with the same mechanics like drugs, alcohol or gambling (Lewis, 2011). Good feeling from longer built achievement is now looking difficult and demotivating. This can lead to resignation to long-term goals. Imagine the people that face this danger as much time as generation Z. It is almost certainly reflected and manifested in their behavior, communication and change in culture. Everyone can modify his image by online tools and in connection with the time spent on social networks it makes a modified image of reality not only for one who modifies but also for one who receives this image communication (Macit, Macit, & Güngör, 2018). The results of research and its outcomes application are evidence of this state.

Dopamine loop is an arising issue connected to behavior and communication change. This state brings to discussion great ethics and security issues connected to usage of social media networks and in development of communication content. Using short-term communication gives a way to make effective communication campaigns but also brings the warning that upcoming generations are passive in dealing with one's own future and life. In addition, they are dealing with a modified image of reality and trust it. Usage of social media and their regulation is the issue to solve for the government and control offices but in the field of research there is new challenge to find the way to activate long-term motivations at least at the level of short term motivations. This would be an interesting topic for further research.

7. Conclusion

Results of the research and its output application is that communication content used for promoting the undergraduate study of information and data science to generation Z is much more effective when short-term topics and goals are included. For the significant decision process is short-term content at least as or more important than the long-term topics. These findings were discussed in the perspective of changing behavior of youths caused by social media and their content's ultimate accessibility which causes addiction on short messages. This phenomenon is called a dopamine loop that works on principle of quick rewards for a brain and is very addictive. By this connection of findings, it reveals the wider importance of this topic especially when it affects such a significant decision as undergraduate study program selection.

Conclusion of these findings is the need for careful acceptance of the state and a careful implementation of short-term communication related principles in information and communication systems and campaigns development. Both communication and user interfaces of information systems may focus on short term rewards to its users and recipients which could be brought by gamification principles for example. Caution must be exercised in balancing the way to efficiency and maintaining a certain ethical and security standard on the other hand. Further research should be focused on finding a way how to deal with long-term communication content for generation Z or in solving associated security and ethical questions.

8. Acknowledgements

9. References

Ahmed, N. (2019). Generation Z's Smartphone and Social Media Usage: A Survey. Journalism and Mass Communication, 9 (3).

Doležalová, J., & Vojtěch, J. (2016). Absolventi středních škol a trh práce: EKONOMICKÁ ODVĚTVÍ. Praha.

Fischer, J. & Vltavská, K. (2016). EUROSTUDENT VI: Základní výsledky šetření postojů a životních podmínek studentů vysokých škol v České republice. Praha.

Hanzalová, K. et al. (2018). Marketingová analýza možností zvýšení počtu zájemců o studium oborů Pojistná matematika a Pojišťovnictví, Fakulta informatiky a statistiky, Vysoká škola ekonomické v Praze.

Koucký, J., & Bartušek, A. (2011). Demografický vývoj a projekce výkonů vysokých škol. Praha: Univerzita Karlova, Pedagogická fakulta, Středisko vzdělávací politiky.

Lewis, M. D. (2011). Dopamine and the Neural "Now". Perspectives on Psychological Science, 6(2), 150-155.

Macit, H. B., Macit, G., & Güngör, O. (2018). Sosyal Medya Bağimliliği Ve Dopamin Odakli Geribildirim Üzerine Bir Araştirma. Mehmet Akif Ersoy Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi, 882-897.

MEDIAN. (2017). MML-TGI v České republice. (n.d.). Retrieved May 03, 2020, from http://showme.median.cz/mml-milenialove/.

Obar, A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue". Telecommunications Policy. 39 (9): 745–750.

Newman, N. et al. (2017). Digital News Report 2017. Retrieved May 03, 2020, from http://www.digitalnewsreport.org/survey/2017/.

Seemiller, C., & Grace, M. (2016). Generation Z goes to college. San Francisco, CA: Jossey-Bass.

Tutgud-Ünal, A. & Deniz, L. (2015). Development of the Social Media Addiction Scale. Online Academic Journal of Information Technology. 21 (6).

COMPARISON OF STUDENT PERFORMANCE IN FULL AND PART-TIME STUDY PROGRAMS BASED ON STATISTICS

Gerhard Jahn

University of Applied Sciences Upper Austria
Gerhard.Jahn@fh-hagenberg.at

Keywords

Part-time study program, drop-out rate, student age, study duration, management data, statistics

Abstract

The drop-out rate among students studying part-time is much higher than full-timers. Clearly, studying while holding down a job does seem to be problematic for many. But why? Obviously, there is pressure on time for part-time students who are also working. But in an attempt to go into the problem in more depth and pinpoint more factors in the higher drop-out rate, this study examines data from a management system that administers student information. As an exemplar, the study focuses on a particular study program, because the curriculum is the same for full and part-time students. And by comparing data, it became obvious that a key factor was the age profile of the student drop-outs. The upshot of our findings is that changes in the organizational aspects of a part-time study program can help to improve this situation.

1. Introduction

Formulating and managing a study program requires a sophisticated approach to the structure and content of the course. Obviously, the aim is to attract well-qualified and motivated students who believe their subject degree will enhance employment opportunities after graduation. But if the course also wants to attract part-time students who have a job and want to continue working while they study for a further qualification, it's essential the course is organised to also cater for them. These students must be confident the course will enable them to balance the change in lifestyle involved in part-time study while continuing to work. Indeed, interested students who are in employment must be assured that they won't have to quit their job if they sign up. And given the high drop-out rate among part-timers generally, it is essential that the right kind of student is enrolled, one who is motivated enough to complete the course even when the going gets tough.

For many years this author has been responsible for the organization of precisely such a part-time study program at the *University of Applied Sciences Upper Austria*. The focus of the program is in information technology, specifically, computer science, hence its title *Software Engineering*, in which graduates obtain a bachelor's degree. All IT courses are offered by our *School of Informatics, Communications, and Media* which is situated on campus in Hagenberg, a village in the northern area of Upper Austria.

Software Engineering is offered as a part-time course of study as well as full-time. The number of student freshers per year is limited to 75, of whom approximately two thirds are full-time students and the remaining third part-time. To qualify for admission, full-time applicants must have graduated

at high-school level. The selection criteria ranking is based on two interviews plus grades at high school. A student's age is not a factor, but the number of older student applicants is practically zero. For part-time students, there is no requirement to be in a job.

The curriculum covers – among other areas – many aspects of programming skills, a foundation in ICT, project management, mathematics, and some non-technical subjects such as social skills and business administration (SE, 2020).

The curriculum is designed to be completed within six semesters. Like many other Universities of Applied Sciences in Austria, we hand out fixed timetables to our students, specifying which lessons they must attend.

Both full and part-time modes of study have identical curricula but – due to the different student profiles – have completely different timetables. Having the same curriculum and in most cases even the same lecturers and learning materials, makes the course a valuable subject for analysis since one can directly compare full and part-time regimes, and in this way, hopefully, discover critical factors militating for or against success.

Details of how the part-time study program differs from its companion have been published by Jahn (2013), the most significant items are:

- Duration of semester is stretched from 15 weeks in the full-time program to 19 weeks for part-time. This reduces the number of lessons per week

- Part-time student contact hours mainly on Fridays and Saturdays plus approximately 25 % of E-Learning.

- Block lectures, which allow final exams to be distributed over the 2nd half of the semester.

2. Methodology

The aim of this paper is to identify factors which influence a student's performance. One usually gets a good impression of a student's state of mind and capabilities during the application interview process. And, obviously, such information gathering is consciously part of that process. Decisions based on such impressions are generally well merited, but it does require a lot of experience – and a bit of luck, too. This paper reflects that process and the management information system used to administer student details like their personal data, their grades, and so on.

We use it (a) to store personal impressions gained, for example, from interviews as valid resource for recall or (b) dismiss it. Obviously, such retained information might help suggest possible changes to the part-time study program that could improve a student's performance.

Raw data from the management information system needs to be cleaned in several aspects, as is discussed in detail below. This is carried out by Microsoft Excel because it is well suited for manual manipulation of data records. Python scripts were written for the data analysis process so that it could be repeatable.

3. Related Work

Student success is an important issue in academic education, and drop-out rates are a key indicator. The reason might be that they are quite easy to observe and record. Staar et al (2018) investigated the reasons for drop-out rates in public administration degree programs, examining the interplay between individual and contextual factors for study success. Interviews with students were their main data source, and the sample was only 147 students, which was quite low.

On the other hand, Belloc et al (2011) used mass data from Italian students, covering more than 100,000 students enrolled on courses between 2001 to 2006. An interesting aspect of their study was that they did not consider drop-out as a binary variable, but as completely separate. Thus, it was either (a) complete withdrawal from academic education or (b) a switch to another study program.

However, this is only observable, when the student's record is accessible even after withdrawal from a study program. Smith & Naylor (2001) examined undergraduate students in the UK. Their conclusion was that important factors for success are (a) pre-academic preparation and (b) social integration at the university.

The *Organisation for Economic Co-operation and Development* yearly publishes statistical data from educational institutions (OECD, 2019). Among many other findings, they report (under *Indicator B5)* how many students completed tertiary education. According to their data, only 39 % of full-time students who enter a bachelor's program graduate within the set duration of the program. The average completion rate after three additional years increased to 67 %.

None of these sources compared part-time students with full-time students within the same study program. The main advantage of using the study program *Software Engineering* is, that we could eliminate non-remediable influence factors like *difficulty* and *amount of time to spend* from the comparison of full and part-time student performance.

Our particular interest in the drop-out rate is occasioned by the limited number of students enrolled each year in Software Engineering: Not selecting the right student applicants results in some lost graduates at the end. In addition, our government funding depends on the number of active students. From the drop-out student's point of view, in most cases the uncompleted study was a waste of time.

4. Data Analysis

4.1. Gathering the Data

The chief source of data for this paper is our in-house student management system which stores data on a relational database. Since direct access to live data is often problematic, duplicates are available for off-line analysis. Access is via ODBC and we use the programming language Python for access. Many sources offer an introduction to Python, some of them even focus on data manipulation and data analysis, such as Grus (2015, pp. 15—36).

In 2003 the study program changed from a diploma to a bachelor's degree, with a consecutive master's program. The bachelor's initially was only on a full-time basis, but in 2004 we added the part-time option. Therefore, data analysis covers this bachelor's program only. Due to internal restrictions, data is only available until summer 2019, but over those 15 years (i.e., since 2004) of parallel operation of both full and part-time programs, the data was enough to give us sufficient insights.

Over these years some adjustments were made, mainly on the curriculum to adapt the contents of lectures to current developments in computer science. All these modifications were applied to both modes of study in parallel, so valid comparisons could be made. In total, we got data from 1,588 students and many samples of grades awarded.

4.2. Cleaning Process

All raw data needs some cleaning which, if not done properly can lead to false outcome conclusions. In our case the most disturbing factor was the existence of more than one record for some students. The reasons for this are:

- Some students began study as "external" students, typically because some of their documentation are missing. Later their status changed to "ordinary". This caused double entries in the management system, which must be merged prior to data analysis.

- When switching from the former diploma program to the bachelor/master's, one class had their first year in the diploma program. Those records need to be merged too.

- Some incoming students stay only for one, or at most two semesters. After this time, they leave the study program without degree. This distorts the drop-out rate and also the duration of study. Such records shall be excluded from further analysis.

The original idea was to automate the whole process of gathering, cleaning and analyzing. The reason for this was so this whole process could periodically be repeated with a simple "button press". To accomplish this, the author started with Python and Pandas (McKinney, 2018) for all parts of this process. But unfortunately, it turned out that the cleaning process needed manual inspection—it could not be done by using just some scripts.

Therefore, good old Excel was used to check for duplicate records of students. Also, the correction of specific data entries was much easier in Excel than in Pandas. Out of almost 1,600 students' records:

- 19 duplicates and 28 incomings were eliminated – some of them even applied to more than two records.

- 33 records were merged with the data from the former diploma study program.

Since the amount of data is far below *big data*, Excel was the correct tool for this operation. However, cleaning the data is not a fully deterministic process. As an example, students sometimes switch during their study from full-time mode to part-time mode and vice versa. This only happens in rare cases, but it is often unclear to which mode of study such a student belongs. This influences the comparison between the organization types – at least to a small degree.

4.3. Condensing Information

Table 1 gives a first insight into the cleaned data.

There are more full-time than part-time students, as defined by study program regulations. Further, more part-time students dropped out of their study than full-time students. This gives a much higher drop-out rate – almost twice as high – which is quite common for part-time study programs in general. In addition, more part-time students are in the state "paused". They take time out of their study, and reasons for this are typically temporary overload due to work-pressures in their job, or their health.

Table 1: Number of regular students in this study program

Students	active	graduate	drop-out	paused	total	drop-out rate
part-time	82	315	240	10	647	37%
full-time	150	537	191	4	882	22%
total	232	852	431	14	1529	28%

Drop-out rate is critical, because of the restricted number of candidates allowed to enrol on the program. For this reason, one is interested in enrolling the "right" applicants, namely those motivated enough to ensure they graduate, and not those prepared to entertain failure. But can statistical analysis help to pick the right students? Among possible reasons why a student may decide to quit studying are:

- Lack of success in exams.

- Lack of time for study, caused by circumstances, like:
  - More workload in their employment.
  - More quality time with family, children.

- Underestimating the effort needed to invest in studying.

The second point, especially, cannot be foreseen in the application process. In general, all points cannot be gleaned from statistical data. One way to discover reasons would be to interview dropping out students, but most are not usually in the mood for this. Another factor that comes up from initial application interviews is the age of candidates. For part-time study, the age of the applicants is higher than those applying to study full-time. This is quite natural, since usually full-time students apply immediately after high school, but part-time applicants are usually already in a job and decide to begin extra study some years afterwards. This observation can easily be verified from our data, which is shown in Table 2.

Table 2: Averaged age of students (in years) at the beginning of their study

Age at entry	all	graduate	drop-out
part-time	28,8	27,4	31,6
full-time	21,5	21,2	22,2

At first sight, it is clear that the part-time group is much older than the other, a difference of more than seven years is very significant. But this table also shows, that older students more often tend to quit their study. While this effect is small in the full-time study program – a difference of one year between graduate students and drop-outs, it's large in the part-time group by more than four years. Looking at this data, one could even say that the drop-out rate is influenced more by the age of the student than by the mode of the study program.

Another point to investigate is the duration that it takes a student to either graduate or quit. This is presented in Table 3.

Table 3: Averaged duration of study (in years)

Duration of study	graduate	drop-out
part-time	3,1	1,5
full-time	2,9	1,6

Concerning this, there is no significant difference between part-time and full-time students.

So far, the presented tables did not consider developments over time. Things may change in the coming years, and we can now look for some trends in the data. For this, the Python/Pandas duo is quite handy, especially for grouping following the split-apply-combine paradigm (Wickham, 2011). The following plots were generated with Matplotlib (McGreggor, 2015) which neatly co-operates with Python.

Figure 1 shows the history of students averaged age over the years. We group by classes, and name a class by the year the students began study.
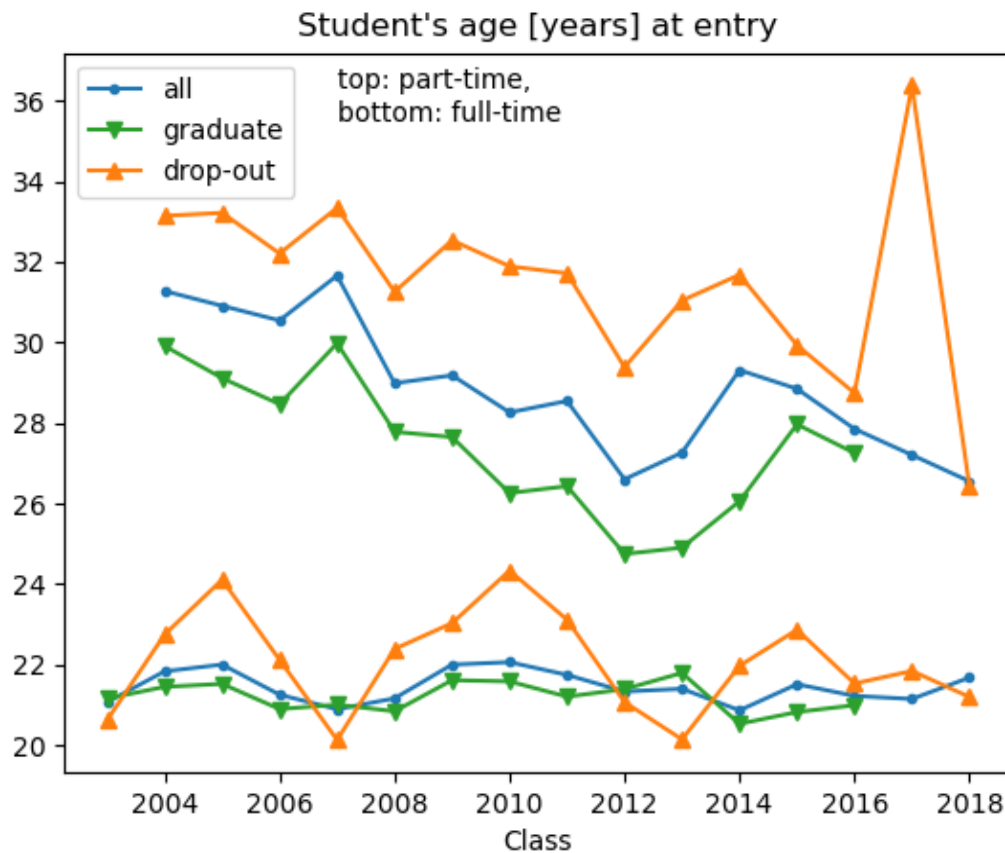
Figure 1: Averaged age of students at their entry, grouped by classes

This figure proves the already observed fact that part-time students are older than their counterparts in full-time study. But it also reveals the age of part-time students enrolling decreases over the years. It also proves that students who quit in the part-time study, are older than those who graduate, no matter what their age of entry into the study program was. Note, that the quite high age of part-time students who quit from the class of 2017 were very few, only six students. Also note that the raw data ends at summer 2019, which is why no graduates from classes 2017 and later exist. The data from full-time students does not show any significant trend. The average age of full-time students who quit is quite stable. This stems from the fact that only very few such students abandon their study.

The history of the drop-out rate is presented in Figure 2. This shows a more or less stable rate in the full-time study program, but a significantly changing one among part-timers. Note, that classes 2017 and later had not yet finished their study. They are "work in progress", which means their drop-out rate will probably grow.
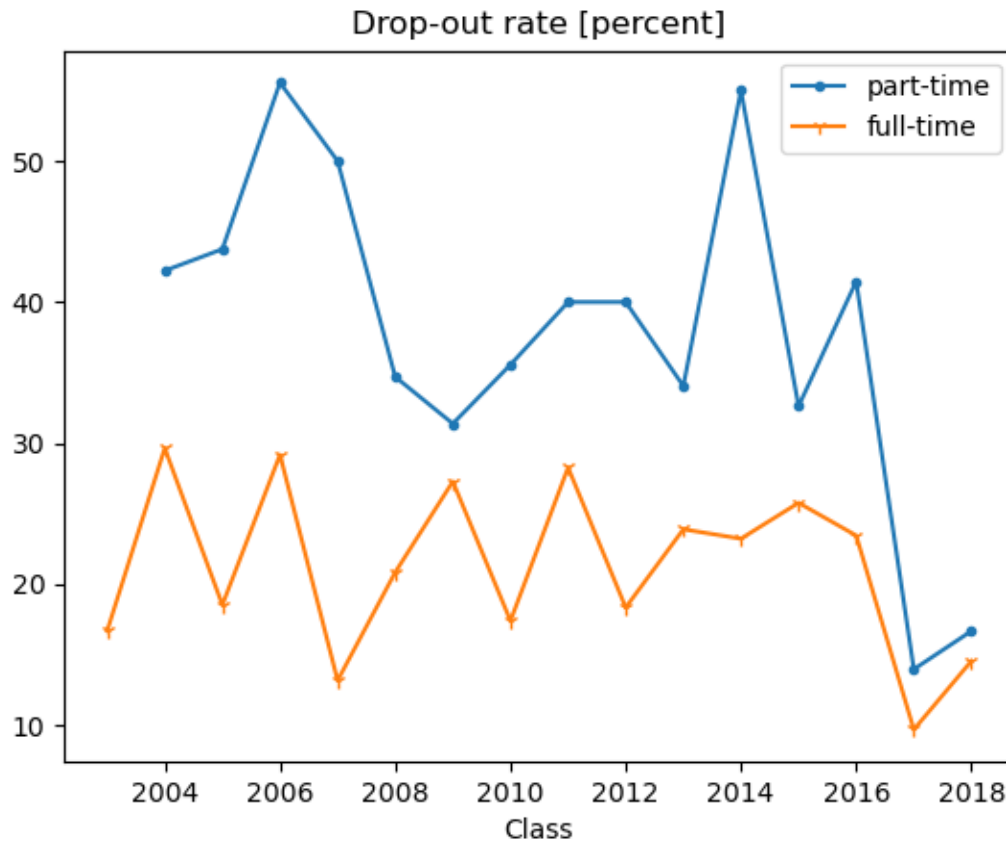
Figure 2: Drop-out rate, grouped by class

5. Conclusion and Future Work

The analysis presented so far shows the assumption that failure rate correlates with student age is correct, to some degree. Although the average age of part-time students decreases over the years, there is only a small decrease in the drop-out rate over the same period. Age, at least, is one of the factors that influences the drop-out rate. Another is the part-time mode itself with the inherent conflict between study, work, and family. The family aspect, especially, counts more among older students since they also tend to have children of their own.

Until now, the drop-out rate of the full-time study program is at 22 % (see Table 1) and quite stable. For the part-time study program, the drop-out rate is at 37 % and slightly trending to decrease. This is almost within the value of 33 % which the OECD gives for full-time study programs (OECD, 2019, Chapter B5).

Prior to this analysis, there was consideration given to stretching the part-time study program by one extra semester and to evenly distribute the courses over those seven semesters. The benefit would be to lower the workload per semester. Up until now, this idea has not been implemented, because the analysis showed that not as many students as we expected were in the state "paused". The drawback of an additional semester would be a less attractive study program for other students who would not welcome this prolongation. Also, we might be precluded from letting full and part-time students

attend the same exams at the same time. This would mean that we would lack some of the comparability for our investigations.

Another interesting effect emerged from the corona crisis during the summer semester of 2020: We had to switch fully to distance learning on all our study programs – which was very well welcomed by our part-time students. What they really appreciated was that they had absolutely no need to travel to Hagenberg to join the lessons. As a result, we plan to continue with more distance learning than we did before the corona crisis.

Although the drop-out rates are currently not at a critical level, an analysis such as this should be repeated approximately every two to three years, in order to identify changing developments. A deeper insight could be gained by also looking at the grades and the changes therein over the years.

A new concept – at least in Austria – are dual systems at an academic level. They combine a study program with the working position of students. One of our departments will start such a dual system in the autumn of 2020 (HSD, 2020). They promise a lesser workload than traditional part-time study programs, and we will be looking closely at this new development.

5. References

Belloc, F., Maruotti, A., & Petrella, L. (2011). How individual characteristics affect university students drop-out: a semiparametric mixed-effects model for an Italian case study, Journal of Applied Statistics, 38:10, 2225-2239, DOI: 10.1080/02664763.2010.545373

Grus, J. (2015). Data Science from Scratch, O'Reilly, Sebastopol, ISBN 978-1-491-90142-7

HSD (2020). Hardware-Software-Design in dualer Studienform (in German), Available at: https://www.fh-ooe.at/campus-hagenberg/studiengaenge/bachelor/hardware-software-design/hsd-dual/duales-studium-hardware-software-design/

Jahn, G. (2013). Organisation und Betrieb eines berufsbegleitenden Studiums am Beispiel des Bachelorstudienganges „Software Engineering" in: Berufsbegleitende Studiengänge als Herausforderung für Curriculumentwicklung und Hochschuldidaktik. Beiträge zum 2. Tag der Lehre an der FH St. Pölten am 12. 4. 2013 (in German), 20—26.

McGreggor, D. M. (2015). Mastering Matplotlib, Packt Publishing, Birmingham, ISBN 978-1-783-98754-2

McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 2nd Edition, O'Reilly, Sebastopol, ISBN 978-1-491-95766-0

OECD (2019), Education at a Glance 2019: OECD Indicators, OECD Publishing, Paris. DOI: 10.1787/f8d7880d-en

SE (2020). Software Engineering–Expertise in software development, with special focus on web or business software, Available at: https://www.fh-ooe.at/en/hagenberg-campus/studiengaenge/bachelor/software-engineering/

Smith, J. P., & Naylor, R. A. (2001). Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students, Journal of the Royal Statistical Society: Series A (Statistics in Society), Vol.164(2), 389–405. DOI: 10.1111/1467-985X.00209

Staar, H., Kania, H., Gurt, J., & Kunert, S. (2018). „Gekommen, um zu bleiben" – eine Analyse des Zusammenhangs zwischen personen- und kontextbezogenen Faktoren und Studienerfolg in Studiengängen der öffentlichen Verwaltung. Gr Interakt Org (in German), 49, 213–229. DOI: 10.1007/s11612-018-0420-z

Wickham, H. (2011). The split-apply-combine strategy for data analysis. Journal of Statistical Software, 40(1), 1—29.

Statement of the Publication Ethics and Publication Malpractice

IDIMT's Publication Ethics and Publication Malpractice Statement is based, in large part, on the guidelines and standards developed by the Committee on Publication Ethics (COPE).

We expect all parties commit to these publication ethics. We do not tolerate plagiarism or other unethical behaviour and will remove any manuscript that does not meet these standards.

The relevant duties and expectations of authors, reviewers, and editors are set out below:

1. Author Responsibilities

Authors must certify that their manuscripts are their original work.

Authors must certify that the manuscript has not previously been published elsewhere.

Authors must certify that the manuscript is not currently being considered for publication elsewhere.

Authors must notify us of any conflicts of interest.

Authors must identify all sources used in the creation of their manuscript.

Authors must report any errors they discover in their manuscript.

2. Reviewer Responsibilities

Reviewers must notify us of any conflicts of interest.

Reviewers must keep information pertaining to the manuscript confidential.

Reviewers must bring to the attention of the Editor-in-Chief any information that may be reason to reject publication of a manuscript.

Reviewers must at any time evaluate manuscripts only for their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors.

Reviewer who feels unqualified to review the research reported in a manuscript or knows that its prompt review will be impossible should notify us and excuse himself from the review process.

3. Editorial Board Responsibilities

The Editorial Board must keep information pertaining to submitted manuscripts confidential.

The Editorial Board must disclose any conflicts of interest.

The Editorial Board must evaluate manuscripts only for their intellectual content.

The Editorial Board is responsible for making publication decisions for submitted manuscripts.

List of Authors

IDIMT-2020
Digitalized Economy, Society and Information Management

28th Interdisciplinary Information Management Talks

Information Management and Interdisciplinarity are keys to foundations of society, economy, and technology, especially in crisis situations. The global accessibility and usability of data offered by Digitalization creates permanent transformations of daily processes and needs and changes everybody's life due to the interdependence of Society, Economy, and Technology. This provides a permanent challenge for our innovative capacity.

The annual IDIMT conference has a 27 year tradition as an international forum for presenting future-oriented and multi-disciplinary research with the emphasis on interdisciplinarity and Information Management.

Based on a double blind review process 43 submitted papers were accepted together with 11 keynotes and an invited papers. The authors come from 12 different countries: Austria, Czech Republic, Estonia, Germany, India, Kazakhstan, Lithuania, Netherlands, Poland, Russia, Slovakia and Ukraine.

The papers have been grouped according to the following themes:
- Innovation, New Business Models and Strategies
- Social Media, Privacy and Ethics
- Digital Economy and Industry 4.0
- Sustainability and Performance Management
- Smart Supply Chain
- Towards 2030 – Digital Transformation in Crisis Management
- Human and Cultural Impacts on Project Success
- Digital Security, Safety, Trust of Smart Systems – for a resilient Society
- People and Processes in Enterprise Software Development
- Cyber Security and Digital Identities