



Leon Strous
Vinton G. Cerf
(Eds.)

Internet of Things

Information Processing
in an Increasingly Connected World

First IFIP International Cross-Domain Conference, IFIPIoT 2018
Held at the 24th IFIP World Computer Congress, WCC 2018
Poznan, Poland, September 18–19, 2018
Revised Selected Papers



Springer Open



Editor-in-Chief

Kai Rannenber, Goethe University Frankfurt, Germany

Editorial Board

TC 1 – Foundations of Computer Science

Jacques Sakarovitch, Télécom ParisTech, France

TC 2 – Software: Theory and Practice

Michael Goedicke, University of Duisburg-Essen, Germany

TC 3 – Education

Arthur Tatnall, Victoria University, Melbourne, Australia

TC 5 – Information Technology Applications

Erich J. Neuhold, University of Vienna, Austria

TC 6 – Communication Systems

Aiko Pras, University of Twente, Enschede, The Netherlands

TC 7 – System Modeling and Optimization

Fredi Tröltzsch, TU Berlin, Germany

TC 8 – Information Systems

Jan Pries-Heje, Roskilde University, Denmark

TC 9 – ICT and Society

David Kreps, University of Salford, Greater Manchester, UK

TC 10 – Computer Systems Technology

Ricardo Reis, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

TC 11 – Security and Privacy Protection in Information Processing Systems

Steven Furnell, Plymouth University, UK

TC 12 – Artificial Intelligence

Ulrich Furbach, University of Koblenz-Landau, Germany

TC 13 – Human-Computer Interaction

Marco Winckler, University of Nice Sophia Antipolis, France

TC 14 – Entertainment Computing

Rainer Malaka, University of Bremen, Germany

IFIP – The International Federation for Information Processing

IFIP was founded in 1960 under the auspices of UNESCO, following the first World Computer Congress held in Paris the previous year. A federation for societies working in information processing, IFIP's aim is two-fold: to support information processing in the countries of its members and to encourage technology transfer to developing nations. As its mission statement clearly states:

IFIP is the global non-profit federation of societies of ICT professionals that aims at achieving a worldwide professional and socially responsible development and application of information and communication technologies.

IFIP is a non-profit-making organization, run almost solely by 2500 volunteers. It operates through a number of technical committees and working groups, which organize events and publications. IFIP's events range from large international open conferences to working conferences and local seminars.

The flagship event is the IFIP World Computer Congress, at which both invited and contributed papers are presented. Contributed papers are rigorously refereed and the rejection rate is high.

As with the Congress, participation in the open conferences is open to all and papers may be invited or submitted. Again, submitted papers are stringently refereed.

The working conferences are structured differently. They are usually run by a working group and attendance is generally smaller and occasionally by invitation only. Their purpose is to create an atmosphere conducive to innovation and development. Refereeing is also rigorous and papers are subjected to extensive group discussion.

Publications arising from IFIP events vary. The papers presented at the IFIP World Computer Congress and at open conferences are published as conference proceedings, while the results of the working conferences are often published as collections of selected and edited papers.

IFIP distinguishes three types of institutional membership: Country Representative Members, Members at Large, and Associate Members. The type of organization that can apply for membership is a wide variety and includes national or international societies of individual computer scientists/ICT professionals, associations or federations of such societies, government institutions/government related organizations, national or international research institutes or consortia, universities, academies of sciences, companies, national or international associations or federations of companies.

More information about this series at <http://www.springer.com/series/6102>

Leon Strous · Vinton G. Cerf (Eds.)

Internet of Things

Information Processing in an Increasingly Connected World

First IFIP International Cross-Domain Conference, IFIPIoT 2018
Held at the 24th IFIP World Computer Congress, WCC 2018
Poznan, Poland, September 18–19, 2018
Revised Selected Papers



Springer Open

Editors

Leon Strous
De Nederlandsche Bank
Amsterdam, The Netherlands

Vinton G. Cerf
Google
Reston, VA, USA



ISSN 1868-4238

ISSN 1868-422X (electronic)

IFIP Advances in Information and Communication Technology

ISBN 978-3-030-15650-3

ISBN 978-3-030-15651-0 (eBook)

<https://doi.org/10.1007/978-3-030-15651-0>

Library of Congress Control Number: 2019934341

© The Editor(s) (if applicable) and The Author(s) 2019. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Like every new technology, the Internet of Things (IoT) offers opportunities for progress and application for beneficial purposes while at the same time it introduces or increases risks and threats. There are many aspects to be considered when talking about IoT. Consequently the IFIP Domain Committee on IoT organized a working conference with a broad scope. In principle, papers on all aspects related to IoT were solicited. This book contains the revised versions of the papers presented at the first IFIP Internet of Things (IoT) conference that took place in Poznan, Poland, during September 18–19, 2018 as part of the IFIP World Computer Congress (WCC) 2018.

The IoT Program Committee consisted of 55 members who considered 24 submissions for the first edition of this conference. Each paper was on average refereed by three reviewers, using the single-blind review principle. In total, 13 papers were selected for presentation resulting in an acceptance rate of 54%. One accepted paper is not included in this book because it was not presented at the conference. The papers were selected on the basis of originality, quality and relevance to the topic.

The papers range from a technology perspective to a business perspective. Topics include hardware, software and management aspects, process innovation, privacy, power consumption, architecture, applications and a few more. In addition to the refereed papers we also have a paper from the invited speaker Kees van der Klauw who challenged the audience by stating that the IoT is hardly about technology. Finally the draft position paper by IFIP on the IoT is included. The paper investigates what choices can or must be made regarding the various aspects of the IoT. This draft was discussed in a panel session and the outcome of the discussion will be included in the final version.

Looking at this wide range of topics makes us realize that we are just at the infancy of the IoT and that a lot of further research and work are needed. We thank the authors, the Program Committee and the participants for their hard work and contributions and look forward to a continued involvement.

The IFIP World Computer Congress (WCC) 2018 had a number of plenary and special sessions scheduled. We are very pleased to present in this book a few contributions out of those sessions. WCC 2018 had four plenary keynote speakers: Wil van der Aalst, Leslie Valiant, Jan Camenish and Shamika Sirimanne. While all four keynote presentations were recorded on video (see www.wcc2018.org and www.ifip.org) Wil van der Aalst also contributed a paper, addressing the question of responsible data science in a dynamic world. A special day at WCC 2018 was the Enigma day with a live demonstration of a message encryption and decryption. A paper in this book describes the history of how three Poznan University students broke the German Enigma Code and shortened World War Two. The third contribution in this book is a summary of workshops on professionalism and frameworks, a “must be” core topic for professional computer societies and associations. And finally, in an open debate, an

emerging question was discussed: Should Artificial Intelligence be more regulated? While this session is also recorded on video, a summary is presented in this book.

We feel that all contributions make the book a rich volume in the IFIP AICT series and we trust that the reader will be inspired by it.

February 2019

Leon Strous
Vinton G. Cerf

IFIP WCC 2018

Information Processing in an Increasingly Connected World: Opportunities and Threats

IFIP WCC 2018 Steering Committee

General Congress Co-chairs

Roman Słowiński	Poznan University of Technology, Poland
Leon Strous	De Nederlandsche Bank, The Netherlands

General Program Co-chairs

Mike Hinchey	Lero-The Irish Software Research Centre, Ireland
Jerzy Nawrocki	Poznan University of Technology, Poland

Publication Chair

Basie von Solms	University of Oxford, UK/University of Johannesburg, South Africa
-----------------	--

General Organizing Chairs

Robert Wrembel (Chair)	Poznan University of Technology, Poland
Andrzej Jaszkiwicz (Co-chair)	Poznan University of Technology, Poland

Members

Joanna Józefowska	Poznan University of Technology, Poland
A Min Tjoa	Vienna University of Technology, Austria

IFIPIoT 2018

Internet of Things: Information Processing in an Increasingly Connected World

IFIPIoT Program Committee

Co-chairs

Leon Strous*	De Nederlandsche Bank, The Netherlands
Vinton Cerf	Google, USA

Members (* are also member of the IFIP Domain Committee IoT)

Jose Abdelnour-Nocera	University of West London, UK
Hamideh Afsarmanesh	University of Amsterdam, The Netherlands
Carmelo Ardito*	Università degli Studi di Bari Aldo Moro, Italy
Ioannis Askoxylakis	FORTH-ICS, Greece
Soumya Banerjee	Birla Institute of Technology, Mesra, India
Ezio Bartocci	Vienna University of Technology, Austria
Juergen Becker	Karlsruhe Institute of Technology, Germany
Samia Bouzeffane	CEDRIC, Conservatoire National des Arts et Métier, France
Luis Camarinho-Matos*	Universidade NOVA de Lisboa, Portugal
Augusto Casaca*	INESC, Portugal
Tibor Cinkler*	Budapest University of Technology and Economics, Hungary
Luc Claessen	University of Hasselt, Belgium
Lucio Davide Spano	University of Cagliari, Italy
Jose Neuman De Souza*	Federal University of Ceara, Brazil
Ibrahim Elfadel	Masdar Institute at Khalifa University of Science and Technology, UAE
Gordon Fletcher*	University of Salford, UK
Miria Grisot	University of Oslo, Norway
Gerhard Hancke	City University of Hong Kong, SAR China
Michael Huebner	Brandenburg University of Technology Cottbus, Germany
Katrin Jonsson	Umeå University, Sweden
Joaquim Jorge	INESC, Portugal
Srinivas Katkoori	University of South Florida, Tampa, USA
Bouabdellah Kechar	University of Oran 1 Ahmed Ben Bella, Algeria
Arianit Kurti	RISE Interactive, Research Institutes of Sweden, Sweden
Maryline Laurent	Telecom SudParis, France
Antonio Mana	University of Malaga, Spain
Tiziana Margaria	University of Limerick, Ireland

Konstantinos Markantonakis	Royal Holloway University of London, UK
Peter Marwedel	Technical University of Dortmund, Germany
Maristella Matera	Politecnico di Milano, Italy
Christina Mörtberg*	Linnaeus University, Sweden
Raja Naeem Akram	ISG-SCC, Royal Holloway University of London, UK
Maciej Ogorzalek	Jagiellonian University, Poland
Fabio Paterno*	CNR-ISTI, Italy
Rasmus Pedersen	Copenhagen Business School, Denmark
Simon Perrault	Yale-NUS College, Singapore
Joachim Posegga	University of Passau, Germany
Ricardo Rabelo	Federal University of Santa Catarina, Brazil
Franz Rammig	University of Paderborn, Germany
Kai Rannenber	Goethe University Frankfurt, Germany
Delphine Reinhardt	Georg-August-Universität Göttingen, Germany
Ricardo Reis*	Universidade Federal do Rio Grande do Sul, Brazil
Carmen Santoro	CNR-ISTI, Italy
Damien Sauveron*	University of Limoges, France
Weiming Shen	National Research Center, Canada
Basie von Solms*	University of Johannesburg, South Africa
Dimitrios Soudris	National Technical University of Athens, Greece
Jean-Yves Tigli	University of Côte d’Azur – UCA, France
Chi Tsui	Hong Kong University of Science and Technology Kowloon, SAR China
Fatih Ugurdag	Ozyegin University Istanbul, Turkey
Eugenio Villar	University of Cantabria, Spain
Janet Wesson	Nelson Mandela University, South Africa
Marco Winckler*	Université Nice Sophia Antipolis, France

Additional Reviewers

Boutheyna Belgacem	University of Passau, Germany
Saifeddine Ben Haj Khalifa	Karlsruhe Institute of Technology, Germany
Akos Grosz	Goethe University Frankfurt, Germany
Javier Hoffmann	Ruhr-University Bochum, Germany
Oswaldo Navarro	Ruhr-University Bochum, Germany
Johannes Pfau	Karlsruhe Institute of Technology, Germany

Contents

WCC 2018 Plenary Contributions: Keynote, Special Sessions

Responsible Data Science in a Dynamic World: The Four Essential Elements of Data Science.	3
<i>Wil M. P. van der Aalst</i>	
How Three Poznan University Students Broke the German Enigma Code and Shortened World War Two	11
<i>Roger G. Johnson</i>	
Professionalism and Frameworks.	21
<i>Moirra de Roche</i>	
Should Artificial Intelligence Be More Regulated? Panel Discussion	28
<i>Leon Strous</i>	

IFIPIoT 2018 Invited Papers: Keynote, Panel Discussion

The Internet of Things is Hardly About Technology	37
<i>Kees van der Klauw</i>	
IoT: Do We Have a Choice? Draft IFIP Position Paper	50
<i>Leon Strous and IFIP Domain Committee on IoT</i>	

IFIPIoT 2018 Refereed Papers

The Outcomes of the Implementation of Internet of Things: A Public Value Perspective	59
<i>Ott Velsberg</i>	
Strategies for Reducing Power Consumption and Increasing Reliability in IoT	76
<i>Ricardo Reis</i>	
An Internet of Things (IoT) Model for Optimising Downtime Management: A Smart Lighting Case Study	89
<i>Brenda Scholtz, Mando Kapeso, and Jean-Paul Van Belle</i>	
IoT Enabled Process Innovation: Exploring Sensor-Based Digital Service Design Through an Information Requirements Framework	105
<i>Niclas Carlén, August Forsman, Jesper Svensson, and Johan Sandberg</i>	

An Internet of Things Based Platform for Real-Time Management of Energy Consumption in Water Resource Recovery Facilities	121
<i>Mário Nunes, Rita Alves, Augusto Casaca, Pedro Póvoa, and José Botelho</i>	
A New Reconfigurable Architecture with Applications to IoT and Mobile Computing	133
<i>Amir Masoud Gharehbaghi, Tomohiro Maruoka, and Masahiro Fujita</i>	
Unexpected Inferences from Sensor Data: A Hidden Privacy Threat in the Internet of Things	147
<i>Jacob Kröger</i>	
Issues in Implementing a Data Integration Platform for Electric Vehicles Using the Internet of Things.	160
<i>Martin Smuts, Brenda Scholtz, and Janet Wesson</i>	
Working with IoT – A Case Study Detailing Workplace Digitalization Through IoT System Adoption	178
<i>Viktor Mähler and Ulrika Holmström Westergren</i>	
Opportunities for the Internet of Things in the Water, Sanitation and Hygiene Domain.	194
<i>Paula Kotzé and Louis Coetzee</i>	
Internet of Things: The Present Status, Future Impacts and Challenges in Nigerian Agriculture	211
<i>Funmilayo O. Bamigboye and Emmanuel O. Ademola</i>	
IoTutor: How Cognitive Computing Can Be Applied to Internet of Things Education	218
<i>Suejb Memeti, Sabri Pllana, Mexhid Ferati, Arianit Kurti, and Ilir Jusufi</i>	
Author Index	235

**WCC 2018 Plenary Contributions:
Keynote, Special Sessions**



Responsible Data Science in a Dynamic World

The Four Essential Elements of Data Science

Wil M. P. van der Aalst^(✉)

Lehrstuhl für Informatik 9, Process and Data Science, RWTH Aachen University,
52056 Aachen, Germany
wvdaalst@pads.rwth-aachen.de
<http://vdaalst.com>

Abstract. Data science is changing our world in many different ways. Data and the associated data science innovations are changing everything: the way we work, the way we move, the way we interact, the way we care, the way we learn, and the way we socialize. As a result, many professions will cease to exist. For example, today's call centers will disappear just like video rental shops disappeared. At the same time, new jobs, products, services, and opportunities emerge. Hence, it is important to understand the essence of data science. This extended abstract discusses the four essential elements of data science: “water” (availability, magnitude, and different forms of data), “fire” (irresponsible uses of data and threats related to fairness, accuracy, confidentiality, and transparency), “wind” (the way data science can be used to improve processes), and “earth” (the need for data science research and education). Next to providing an original view on data science, the abstract also highlights important next steps to ensure that data will not just change, but also improve our world.

Keywords: Data science · Responsible data science · Process mining · Big data

1 Data Science

This extended abstract is based on a keynote given at the IFIP World Computer Congress (WCC 2018) on 18 September 2018, in Poznan, Poland. The main theme of WCC 2018 was “Information Processing in an Increasingly Connected World: Opportunities and Threats”. Data science is the main driver for the changes that create these opportunities and threats. Recent reports [6, 7] indicate that many jobs will cease to exist because of advances in machine learning, artificial intelligence, robotics, and other forms of smart automation. These advances are only possible because of both the availability of data and progress in data science.

It is not easy to define data science. The data science pipeline shown in Fig. 1 illustrates the breadth of the discipline. The “infrastructure” part of the pipeline

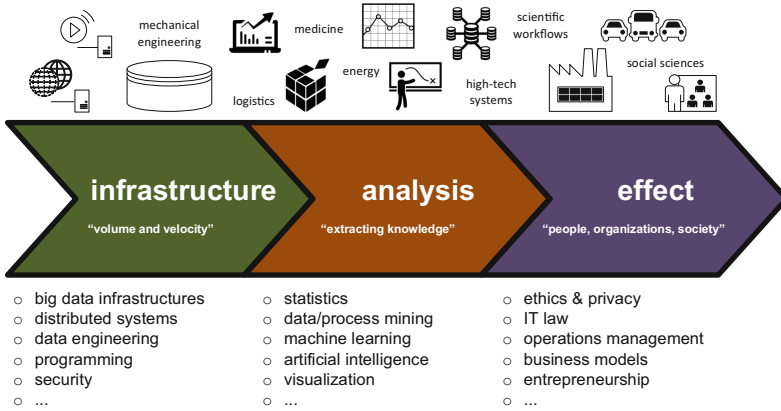


Fig. 1. The data science pipeline showing that different capabilities are needed to turn data into value.

is concerned with the huge volume and incredible velocity of data. Hence, the primary focus is on making things scalable and instant. The “analysis” part of the pipeline is concerned with extracting knowledge. This is about providing answers to known and unknown unknowns.¹ The “effect” part of the pipeline is concerned the impact of data science on people, organizations, and society. Here legal, ethical, and financial aspects come into play.

The uptake of the Internet of Things (IoT) illustrates the pivotal role of data science. More and more devices (light bulbs, clothes, refrigerators, containers, bicycles, etc.) are connected to the internet and produce data. These devices are becoming “smart” by learning from the data collected. The Internet of Things (IoT) depends on the whole data science pipeline shown in Fig. 1. We are (or will be) surrounded by smart devices collecting data and the impact of this cannot be overestimated.

In the remainder, we define the four essential elements of data science. As metaphor we use the classical four elements: “water”, “fire”, “wind”, and “earth”. According to the Empedocles, a Greek pre-Socratic philosopher who lived in Sicily in the fifth century B.C., all matter is comprised of these four elements. Other ancient cultures had similar lists, sometimes also composed of more elements (e.g., earth, water, air, fire, and aether) that tried to explain nature and complexity of all matter in terms of simpler substances. Today, we know that this is not the case. However, for data science, we are still in the phase where we are looking for the essential elements. This paper uses “water” as a placeholder for the availability of different forms of data, “fire” as a placeholder for irresponsible uses of data (e.g., threats to fairness, accuracy, confidentiality,

¹ “There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don’t know we don’t know.” (Donald Rumsfeld, February 12, 2002).

and transparency), “wind” as a placeholder for the way that data science can be used to improve processes, and “earth” as a placeholder for education and research (i.e., the base of data science) underpinning all of this. These four essential elements are discussed in the remaining sections.

2 The “Water” of Data Science

The first essential element of data science (“water”) is the data itself. The exponential growth of data is evident. Figure 2 (inspired by the analysis in [9]) shows the rapid developments in terms of *costs* (things are getting exponentially cheaper), *speed* (things are going exponentially faster), and *miniaturization* (things are getting exponentially smaller). This is not limited to *processing* (i.e., CPU and GPU processors), but also applies to *storage* and *communication*. Consider for example the costs of storage. To store one megabyte (MB) of data in the sixties one would need to pay one million euros. Today, one can buy a 10TB harddisk for less than 300 euro, i.e., 0.00003 cents per MB. Another example is the bandwidth efficiency, also called spectral efficiency, which refers to the information rate that can be transmitted over a given bandwidth. It is the net bitrate (useful information rate excluding error-correcting codes) or maximum throughput divided by the bandwidth in hertz of a communication channel or a data link. The spectacular progress of our data handling capabilities illustrated by Fig. 2, explains why data science has become one of the key concerns in any organization. In the sixties, we only had a few “drops of data” whereas today we are facing a “tsunami of data” flooding our society.

Clearly, data science has its roots in statistics, a discipline that developed over four centuries [1]. John Graunt (1620–1674) started to study London’s death records around 1660. Based on this he was able to predict the life expectancy of a person at a particular age. Francis Galton (1822–1911) introduced statistical concepts like regression and correlation at the end of the 19th century. Although data science can be seen as a continuation of statistics, the majority of statisticians did not contribute much to recent progress in data science. Most statisticians focused on theoretical results rather than real-world analysis problems. The computational aspects, which are critical for larger data sets, are typically ignored by statisticians. The focus is on generative modeling rather than prediction and dealing with practical challenges related to data quality and size. When the data mining community realized major breakthroughs in the discovery of patterns and relationships (e.g., efficiently learning decision trees and association rules), most statisticians referred to these discovery practices as “data fishing”, “data snooping”, and “data dredging” to express their dismay [1, 4, 10].

Put differently; most statisticians were focused on techniques to make reliable statements given a few “drops of data”. Such viewpoints turned out to be less effective when dealing with “tsunamis of data”.

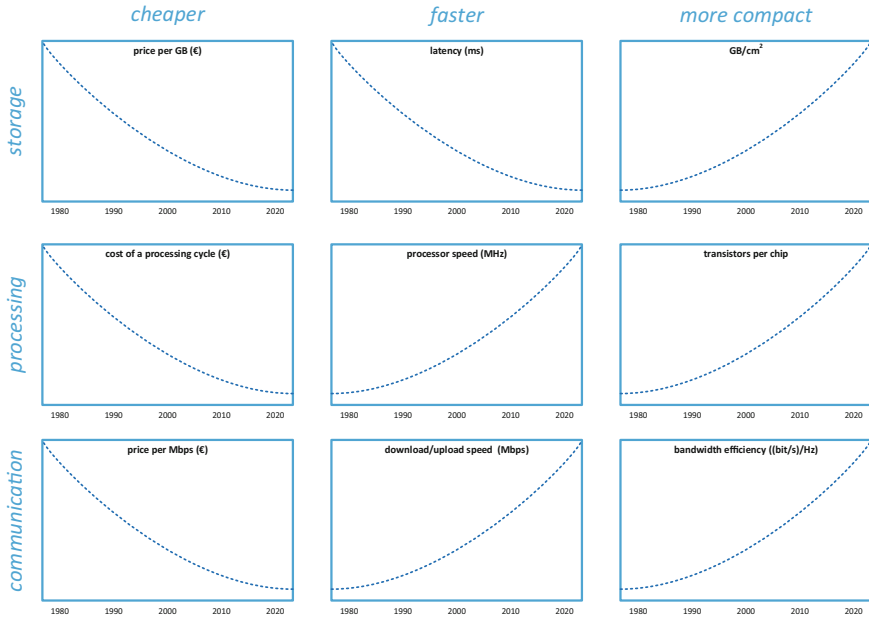


Fig. 2. Moore’s law predicts an exponential growth of the number of transistors per chip. This can be generalized to storage and transition and also applies to costs and speed.

3 The “Fire” of Data Science

The second essential element of data science (“fire”) refers to the dangers of using data in an irresponsible way. Data abundance combined with powerful data science techniques has the potential to dramatically improve our lives by enabling new services and products, while improving their efficiency and quality. Many of today’s scientific discoveries (e.g., in health) are already fueled by developments in statistics, mining, machine learning, artificial intelligence, databases, and visualization. At the same time, there are also great concerns about the use of data. Increasingly, customers, patients, and other stakeholders are concerned about irresponsible data use. Automated data decisions may be unfair or non-transparent. Confidential data may be shared unintentionally or abused by third parties.

From 2015 until 2017, the author led the *Responsible Data Science* (RDS) initiative where the strongest Dutch data science groups joined forces to address problems related to *fairness*, *accuracy*, *confidentiality*, and *transparency* (www.responsibledatascience.org). The goal of RDS is to show that data science techniques, infrastructures and approaches can be made responsible by design. *Responsible Data Science* (RDS) revolves around four main challenges:

- *Data science without prejudice* - How to avoid unfair conclusions even if they are true?

- *Data science without guesswork* - How to answer questions with a guaranteed level of accuracy?
- *Data science that ensures confidentiality* - How to answer questions without revealing secrets?
- *Data science that provides transparency* - How to clarify answers such that they become indisputable?

The term *green data science* was introduced for cutting-edge solutions that enable individuals, organizations and society to benefit from widespread data availability while ensuring *Fairness, Accuracy, Confidentiality, and Transparency* (FACT) [2].

Naïvely one could think that “fire” can be controlled by “water”, however this is not the case. When considering RDS, it is better to consider data as “oil” rather than “water”. It needs to be controlled and stored carefully.

There is a need for new and positive data science techniques that are responsible (i.e., “green”) by design. This cannot be solved by stricter laws. Using the metaphor of “green energy”: We should not be against the use of energy (“data”), but address the pollution caused by traditional engines. Fortunately, there are plenty of ideas to make data science green. For example, discrimination-aware data mining [8] can be used to ensure fairness and polymorphic encryption can be used to ensure confidentiality.

4 The “Wind” of Data Science

The third essential element of data science (“wind”) is concerned with the way data and processes interact. Storing and processing data is not a goal in itself. Data are there to support processes. The campaign “The best run companies run SAP” illustrates that the purpose of information systems is to ensure that processes run well. Data science can help organizations to be more effective, to provide a better service, to deliver faster, and to do all of this at lower costs. This applies to logistics, production, transport, healthcare, banking, insurance, and government. This also applies to individuals. Data science will increasingly support our personal workflows and take over tasks, or at least support them. Data (“water”) can be used to manage and support processes (“wind”) through the use of data science technologies.

An emerging technology linking “water” and “wind” is *process mining* [1]. Process mining bridges the gap between traditional model-based process analysis (e.g., simulation and other business process management techniques) and data-centric analysis techniques such as machine learning and data mining. Process mining seeks the confrontation between event data (i.e., observed behavior) and process models (hand-made or discovered automatically) [1]. The process-mining spectrum is broad and includes techniques for process discovery, conformance checking, prediction, and bottleneck analysis. These techniques tend to be very different from mainstream data mining and machine learning techniques which are typically not process-centric.

Consider for example the topic of *Robotic Process Automation* (RPA). RPA is an umbrella term for tools that operate on the user interface of other computer systems in the way a human would do. RPA aims to replace people by automation done in an “outside-in” manner [3]. This differs from the classical “inside-out” approach to improve information systems. Unlike traditional workflow technology, the information system remains unchanged. The robots are replacing humans while leaving the back-end systems intact. RPA is a way to support processes in a more cost-effective manner. However, this requires learning what humans do by observing them. Data science approaches like process mining can be used to learn the behavior of people doing routine tasks. After the desired behavior has been “played in”, it can be “played out” to handle new cases in an intelligent manner.

RPA illustrates that data science will lead to new trade-offs between what humans do and what robots do [6, 7]. These trade-offs are interesting: How to distribute work between given breakthroughs in data science? Obviously, the question needs to take the “fire” dimension into account.

5 The “Earth” of Data Science

The fourth essential element of data science (“earth”) is concerned with the foundations of a data-driven society: *education* and *research*. Education (in every sense of the word) is one of the fundamental factors in the development of data science. Data science education is needed at any level. People need to be aware of the way algorithms make decisions that may influence their lives. Privacy discussions reveal the ignorance of policy makers and end users. Moreover, to remain competitive, countries should invest in data science capabilities. This can only be realized through education. Data science research plays a similar role. On the one hand, it is key for our education. On the other hand, research is needed to address the many technological and societal challenges (e.g., ensuring fairness, accuracy, confidentiality, and transparency).

Currently, eight of the world’s ten biggest companies, as measured by market capitalization, are American: Apple, Alphabet (incl. Google), Microsoft, Amazon, Berkshire Hathaway, Facebook, JPMorgan Chase, and Bank of America.² The two remaining companies are Chinese: Alibaba and Tencent Holdings. This shows the dominance of a few countries due to investments in IT. Most of the companies are relatively new and emerged through the smart use of data. Amazon and Alibaba are dominating the way we buy products. Google is controlling the way we search. Facebook is controlling the way we socialize. Apple, Alphabet, and Microsoft are controlling the platforms we use (iOS, Android, and Windows). Consider for example Facebook. On the one hand, many people are expressing concerns about the use of data. On the other hand, Facebook has over 2 billion monthly active users that provide personal information in order to use social media. One of the problems of data science is that due to economies

² Based on market capitalization data by Bloomberg on 31 March 2018.

of scale “the winner takes it all”. This may also apply to education, e.g., on Coursera a few US universities are dominating data science education.

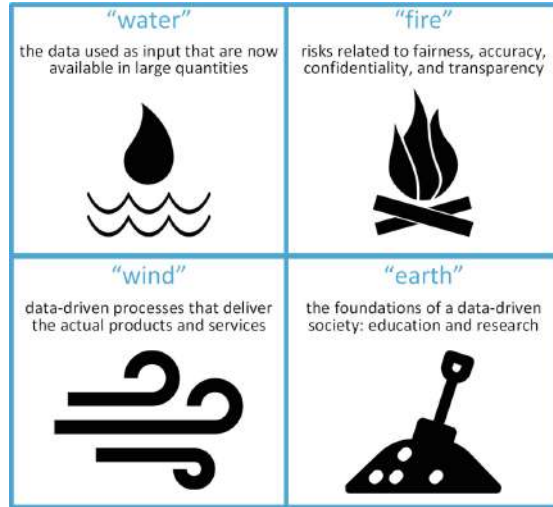


Fig. 3. The “water”, “fire”, “wind”, and “earth” of data science.

Data science literacy and major public investments are needed to address these concerns. This cannot be left to “the market” or solved through half-hearted legislation like the European General Data Protection Regulation (GDPR) [5].

6 Epilogue

This extended abstract aimed to present some of the key messages of the keynote presentation for the IFIP World Computer Congress (WCC 2018). It stresses the importance of data science for people, organizations, and society. Just like computer science emerged as a new discipline from mathematics in the early eighties, we can now witness that the data science discipline is emerging from computer science, statistics, and social sciences.

In this paper, we discussed the four essential elements of data science (see Fig. 3): “water” (availability, magnitude, and different forms of data), “fire” (irresponsible uses of data and threats related to fairness, accuracy, confidentiality, and transparency), “wind” (the way data science can be used to improve processes), and “earth” (the need for data science research and education). By presenting data science in this manner, we hope to get more attention for process-centric forms of data science (e.g., process mining), responsible data science, data science education, and data science research. The dominance of a few companies and countries when it comes to data science is undesirable and requires

the attention of politicians and policymakers. The IFIP could and should play an active role in this discussion.

References

1. van der Aalst, W.M.P.: *Process Mining: Data Science in Action*. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49851-4_1
2. van der Aalst, W.M.P.: Responsible data science: using event data in a “People Friendly” manner. In: Hammoudi, S., Maciaszek, L.A., Missikoff, M.M., Camp, O., Cordeiro, J. (eds.) *ICEIS 2016. LNBIP*, vol. 291, pp. 3–28. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62386-3_1
3. van der Aalst, W.M.P., Bichler, M., Heinzl, A.: Robotic process automation. *Bus. Inf. Syst. Eng.* **60**(4), 269–272 (2018)
4. Breiman, L.: Statistical modeling: the two cultures. *Stat. Sci.* **16**(3), 199–231 (2001)
5. European Commission: Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation). 9565/15, 2012/0011 (COD), June 2015
6. Frey, C.B., Osborne, M.A.: The future of employment: how susceptible are jobs to computerisation? *Technol. Forecast. Soc. Change* **114**, 254–280 (2017)
7. Hawksorth, J., Berriman, R., Goel, S.: Will robots really steal our jobs? An international analysis of the potential long term impact of automation. Technical report, PricewaterhouseCoopers (2018)
8. Pedreshi, D., Ruggieri, S., Turini, F.: Discrimination-aware data mining. In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 560–568. ACM (2008)
9. Brennenraedts, R., Vankan, A., te Velde R., Minne, B., Veldkamp, J., Kaashoek, B.: The impact of ICT on the Dutch economy. Technical report, Dialogic (2014)
10. Tukey, J.W.: The future of data analysis. *Ann. Math. Stat.* **33**(1), 1–67 (1962)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





How Three Poznan University Students Broke the German Enigma Code and Shortened World War Two

Roger G. Johnson^(✉)

School of Computer Science, Birkbeck University of London,
Malet Street, London WC1E 7HX, UK
rgj@dcs.bbk.ac.uk

Abstract. The story of the Allied breaking of the German Enigma codes in World War 2 was first published in the 1970s. Even now many of the details, especially concerning the critical work in the 1930s undertaken by gifted and dedicated Polish codebreakers remains largely unknown. Their work is credited with saving the Allies several years work and so shortening the war and saving thousands of lives. The holding of the IFIP World Computer Congress in Poznan, home of the Polish codebreakers, gave an opportunity for their work to be highlighted to an international audience. Talks covering the work of the Polish, British and French codebreakers were given and webcast worldwide. In addition, a encoded Enigma message was sent at the start of the day from Poznan to Bletchley Park in the UK where the volunteers of the Bombe team at The National Museum of Computing successfully confirmed their breaking of the message at the start of the afternoon session.

Keywords: Enigma · Code breaking · World War II · Marian Rejewski · Jerzy Rozycki · Henryk Zygalski · Turing-Welchman Bombe

1 Background

In 1945 General Dwight D Eisenhower (Allied Supreme Commander Europe) wrote to General Stewart Menzies (Head of Bletchley Park in the UK saying that the successful reading of German messages had

“saved thousands of British and American lives and, in no small way, contributed to the speed with which the enemy was routed and eventually forced to surrender”

The story of how the Allied forces broke the German Enigma code during World War 2 has been told many times in recent years usually from a variety of perspectives mostly linked to Bletchley Park in the UK. The critical contribution of the Polish codebreakers remains little known outside Poland and only a limited number of books and papers have been published about their work. The Polish codebreakers repeatedly broke the Enigma code as its security features were steadily enhanced throughout the 1930s. The result was that as war was about to break out in 1939 the Poles were able to give working replica Enigma machines to their French and British allies and to explain how they had successfully broken the German Enigma messages up to that time.

Without this dramatic gesture it is very unlikely that the British and French would have been able to develop the codebreaking techniques which enabled the British to read German Enigma traffic throughout most of World War 2 at Bletchley Park and also the French until late 1942 at Bletchley Park's French equivalent.

This paper summarises the story of the critical Polish contribution and how it was built on by the French and British, most notably by the British mechanisation of the most time-consuming part of finding the key each day by the building of machines which were named Bombes. The Polish role is well documented in two books, Kozaczuk (1984) first published in Polish in 1979 which focussed primarily on the codebreaking and very recently in (Turing 2018) which recounts the codebreaking exploits but also the lives of the codebreakers during and after this tempestuous period.

The holding of the IFIP World Congress in Poznan in Poland provided an ideal opportunity for IFIP WG 9.7 on the History of Computing to celebrate the work of three talented and heroic Polish mathematics students from the University of Poznan, Marian Rejewski, Jerzy Rozycki and Henryk Zygalski, who trained in Poznan to become codebreakers and whose work ultimately led to the significant shortening of World War 2.

The author is a member of IFIP WG 9.7 and is also the Secretary of the Turing Welchman Bombe Rebuild Trust (TWBRT) which owns the replica Bombe completed in 2007 and is demonstrated every week at The National Museum of Computing (TNMoC) housed in Block H of Bletchley Park in the UK. He arranged for the TWBRT Bombe team to hold one of its occasional roadshow events in which an Enigma message is sent from a remote location to the Bombe Team at TNMoC who then attempt to break the code and send back confirmation of the message being successfully read.

2 Enigma Machine

The origins of the Enigma coding machine were with a commercial coding machine built by a German electrical engineer named Arthur Scherbius. He obtained several patents for his machines starting in 1918. The device evolved into a portable device about the size of a typewriter powered by batteries. Having initially failed to interest the German armed forces in his machine he sold them as commercial coding machines for use by financial institutions such as banks to protect commercially sensitive information being sent by telegraph and other devices. Turing (2018) records that in 1926 both the British and Polish authorities had obtained commercial examples to study while he also notes that, in the same year, German Navy signals using an Enigma machine are noted for the first time.

Following the largely static army operations of the First World War, military strategists developed ideas for future mobile land warfare. However, a critical issue would be to create effective communications for command and control of relatively small frontline military units. In addition, naval commanders, especially with a growing force of submarines, needed secure two-way communications to maintain contact with their forces. What was required was a secure coding machine which, given its presence close to the frontline in mobile warfare, could sooner or later be captured by the enemy

without compromising the security of the communications network. This need for portable, secure communications potentially across large distances, was the capability the Enigma machine provided. Figure 1 shows a German army Enigma machine.



Fig. 1. Three wheel Enigma machine

The key features of the Enigma machine were a conventional German keyboard and above it lamps which light each time a key is pressed with the enciphered character corresponding to the key pressed. Above the lamps are the three rotor wheels. Each time a key is depressed the righthand rotor advances one step and after one revolution the adjacent rotor advances one step and similarly with the leftmost rotor. Each rotor has the letters of the alphabet around the rim and every letter is wired to another letter elsewhere on the rotor. Thus a letter “A” typed on the keyboard may emerge from the first rotor as “K” and so on through the other two rotors. The electric current then reaches the plugboard on the front of the machine where 20 of the 26 letters are again wired up in pairs after which it returns through the wheels until it lights up a lamp on the machine. The Army Enigma machine ultimately had five rotors and on any day three would be used in a predefined arrangement of rotors. The result of all these different combinations is to produce over 150 million alternatives. A very comprehensive account of the evolution of the Enigma coding machine is provided in (Perera 2010).

It was this extraordinary number of combinations which several times later in the war led the Germans to conclude, in the face of circumstantial evidence to the contrary such as dramatic increases in submarine losses after the Allied breaking of the naval Enigma, that the Enigma machine had in fact not been broken but that there was an

alternative explanation, such as espionage or allied technological advances, for significant German setbacks.

It is worth noting that other military powers, including Britain, also adopted coding machines which made use of rotors. It was fundamentally a good approach to automated enciphering of messages in an electro-mechanical era.

3 Poland and Germany

The inter-war Polish state was a creation of the Versailles Peace treaty. It was situated between Germany and Russia and the Polish authorities trusted neither. In the turmoil following the Russian revolution, the Polish government regarded the German state of the later 1920s as a bigger potential threat than the Soviet Union. Also Soviet codes were still using First World war techniques and so liable to successful attack.

Initial attempts to break German Enigma messages using the commercially available Enigma machine failed. Obviously the machine had been modified. Any attack on the machine would need trained cryptographers and so a special course was run at the University of Poznan which was in a part of Poland which had formerly been part of Germany and hence had many fluent German speakers. In 1929 20 students were recruited to the course. Further attempts at breaking into Enigma still yielded nothing until in 1932 the French recruited a spy, Hans-Thilo Schmidt, who worked as a civilian in the German Army's cryptography unit. To fund an extravagant life style he needed money and proceeded to sell large numbers of photographs of secret files relating to the work of the cryptographic unit to the French.

Unfortunately without a German military Enigma machine the French realised that the photographs of the operating instructions were of no immediate value. The British when offered the photographs came to the same conclusion. The French then approached the Poles who expressed more interest but asked for more information. Gradually through the first half of 1932 the French obtained more and more material from Hans-Thilo Schmidt until finally, in August 1932, they obtained an encrypted message together with the original text. With the other secret material already obtained, it now appeared that it might be possible to reverse engineer the Enigma machine, in particular the wiring of the rotors.

The first recruit to the Polish Cypher Bureau from the Poznan course was Marian Rejewski. Initially he worked on Enigma in the evening after his colleagues in the Cypher Bureau had gone home. Later on it became a full time but still secret project. Month by month he gradually worked out the wiring inside the machine. While for their part the French continued to supply more secret intelligence from Hans-Thilo Schmidt. The final problem to be overcome, once the wiring of the Enigma machine had been worked out was to determine a way to find which rotors were being used, in what order they had been placed into the machine and the starting position for each of the rotors. Marian Rejewski noticed that each message began with the starting position sent twice.



Fig. 2. Marian Rejewski, Jerzy Rozycki, Henryk Zygalski

Marian Rejewski was now joined by two more graduates of the Poznan course. They were Jerzy Rozycki and Henryk Zygalski (Fig. 2). From the stolen operating instructions they knew that the Germans sent the starting position twice at the start of each message and they realised that during the encipherment almost certainly only the righthand wheel turned while the others remained stationery. Studying the patterns enabled them to devise simple lookup methods to find the arrangement of the rotors and also some of the plugboard settings. Further they realised that what they needed was a working copy of an Enigma machine. Starting with an old commercial machine as a model, the Poles constructed in utmost secrecy a small number of machines functionally the same as the then current German Enigma machine complete with correctly wired rotors and a plugboard.

The procedures used by the Germans continued to evolve. Gradually rotor orders were changed more frequently until in October 1936 they were changed daily. Sloppy operating practices were eliminated and more cables were used on the plugboard. A major change took place as war clouds gathered in 1938 when the Germans introduced two new rotors, making five in total, and changed their operating procedure to use a different initial wheel position for each message. Each time the Poles responded with new techniques to re-establish the setup of the machine so that messages could be successfully read.

4 Sharing with Britain and France

By 1938 both the British and French cryptographers had looked at approaches to breaking the Enigma messages but had made little progress with the latest German versions of the machine. They had only succeeded in breaking into simpler versions of Enigma used in the Spanish Civil War and also the less advanced Italian system.

The Munich crisis of 1938 caused both to examine their readiness for war which led to a substantial exchange of information about Enigma. The French knew from the intelligence they had supplied to the Poles that the Poles had probably made some progress but the British appear to have been unaware of the possible significance of the Polish work. However, the major changes by the German in late 1938 had stretched the

Polish resources close to breaking point. Their productivity in breaking into Enigma had dwindled dramatically.

In December 1938 the French proposed holding a three way conference in Paris between France, Britain and Poland at which the French hoped to find out what progress each had made. The meeting, held in January 1939, went badly with each party revealing only very limited amounts of information. However, it was clear to each party that the others were serious in their commitment to break into Enigma and so contacts were maintained through the spring and summer of 1939.

The next meeting was to be truly momentous but neither the British or French knew in advance. At the end of June 1939 the Poles, knowing through Enigma and other intelligence that Germany was preparing to invade Poland, invited the British and the French to Warsaw for a meeting. Thus it was in late July 1939 Alastair Denniston, Head of Bletchley Park and Dilly Knox, Britain's leading cryptographer and their principal expert on Enigma travelled across Nazi Germany by train to Poland. The French were represented by Gustav Bertrand, Denniston's opposite number and his deputy, Henri Braquenie. The Poles sent their trio of Rejewski, Rozycki and Zygalski together with their boss, Maksymilian Ciezki.

On the day following their arrival they were driven to the Poles' secret intelligence HQ at Pyry on the outskirts of Warsaw. To the amazement of the French and British the Poles announced almost immediately that they had broken Enigma some years earlier. The Poles showed them a variety of devices which they used to help determine each day's Enigma settings. Discussions continued next day as the Poles revealed more of their methods for breaking the code. However, without doubt, the highpoint was the offer by the Poles to donate to both the French and the British one of their precious working replica Enigma machines. The two machines left Poland by diplomatic bag for Paris and so, probably unnoticed by fellow travellers, Stewart Menzies, the Deputy Head of the British Secret Intelligence Service greeted Gustav Bertrand, the Head of the French Codebreakers as he arrived at Victoria Station in August 1939 with a large wooden box containing the priceless Enigma machine donated to the British.

At this point Alan Turing enters the story. He had been working part time on the Enigma problem at Cambridge since 1938 but had not made much progress. Following the Pyry meeting, Knox had shared with Turing all the information that the Poles had provided, including their mechanical devices for finding the key of the day. Very rapidly Turing conceived of an electro-mechanical machine to search for feasible solutions for the rotor starting positions based on a technique of guessing what the often stylised clear text of the German Enigma message might be. This specification for a machine was handed to BTM, the UK's leading punched card equipment manufacturer who were closely tied to IBM based in the USA, to turn into a physical reality.

A clear and full account of what became known as the Turing Welchman Bombe and how it was used is given in (Turing [2014](#)).

5 After the Polish Invasion

On September 1st 1939 Germany invaded Poland and by the end of the month Polish resistance had collapsed. Poland was divided into three with large parts being assimilated by Germany in the west and the Soviet Union in the east with a small central area under the control of the Polish General Government. The Poles had planned for an invasion and destroyed evidence of their Enigma codebreaking work. It was vital that the codebreakers got away and so travelling by train and lorry they fled to Romania where they went first to the British Embassy who did not appreciate their significance and asked them to return the following day after the staff had contacted London. However, if they had been caught by the Romanian secret police they would probably be handed over to the Gestapo. Consequently, the Poles moved on immediately to the French Embassy who recognised their links with the French Secret Service and assisted them to reach France where they were met at the border by a representative sent by Gustav Bertrand. Knox and Denniston were not amused to find that the French had now got all the key Polish Enigma experts.

There followed a period of cooperation between Bletchley Park and the French codebreakers now established in the Chateau de Vignolles near Paris. The two groups were linked by a secure landline and from early 1940 there were daily races to find the Enigma key of the day. However, this period was not to last long. Early in May 1940 the German Army attacked the French and British forces in the West and on June 25th an armistice was signed between Germany and France. This divided France into two main areas – Occupied France in the north and “Free France” in the south with its government based in the small spa town of Vichy. From there, the Vichy government ran both Vichy France and also the whole of the worldwide French colonial empire.

In the anticipation that there might be an underground resistance movement within Vichy France, the armistice permitted the Vichy government to maintain a small codebreaking capability to track them down although they were expressly forbidden from intercepting German messages. Bertrand’s group, including the Poles, moved to form this group now relocated to a small chateau outside Uzès near Nîmes in southern France. The group now continued to intercept message traffic including German Enigma messages. Intelligence obtained, depending on its contents, could be passed to the Vichy authorities or to other groups. Bertrand’s group built up a network of links across north Africa and Portugal supplying intelligence directly and through intermediaries to the British as well as De Gaulle’s Free French and the Polish Government in exile in London and received equipment, finance and other benefits in exchange.

Assorted codebreakers travelled between the chateau at Uzès and north Africa to meet with other units working there. One of these trips ended in disaster when in January 1942 Jerzy Rozycki was drowned, when the ship on which travelling back to France from Algiers foundered in heavy seas with a substantial loss of life.

North Africa was in a very fluid state with many loose loyalties. In some places, such as Tangier, which had an international zone, officials as well as agents from many of the warring powers rubbed shoulders throughout the conflict. Fascinating insights into this period are to be found in (Pidgeon 2008) which includes material on North Africa.

In November 1942, German and Italian forces took over Vichy France. The German authorities and their Vichy collaborators were closing in on the radio transmissions from the chateau. It was decided that the Poles should leave. The British concluded that the Poles were too numerous to be flown out. The other alternatives were to attempt an evacuation by sea, or overland via Switzerland or Spain. However the route into Switzerland was now effectively closed. Attempts to evacuate by sea proved too dangerous. Consequently in early January 1943 groups of Polish codebreakers began to travel across France towards the Pyrenees and the Spanish border. Marian Rejewski and Henryk Zygalski managed with some difficulty to cross the Spanish border together. In common with most undocumented entrants into Spain they were jailed by the Spanish authorities. However, as the German and Italian armies suffered reverses the attitude of the Spanish authorities softened. Finally, starting in April 1943 the prison camps were gradually emptied. Marian Rejewski and Henryk Zygalski were finally released and by stages travelled via Portugal and Gibraltar to the UK. Having regained their freedom they were once again part of the Polish armed forces. They were attached to a team based near Hemel Hempstead which worked on Russian codes for the remainder of the war.

When peace returned to Europe in May 1945, Marian Rejewski and Henryk Zygalski both faced a difficult choice, whether to return to Poland or to find a new home. Marian Rejewski had a wife and two small children in Poland and so he decided to return to his homeland. Returnees were often regarded with suspicion by the new communist authorities in Poland. Although his career as an accountant was interfered with by the authorities due to suspicions about his wartime work he survived to be honoured by Poland prior to his death in 1980 for his services to the defeat of Germany as the Polish political environment evolved. Henryk Zygalski in contrast had met a British girl during his wartime work in the UK. He became a British citizen and settled down to an academic career in the UK ultimately as a member of staff of the Mathematics Department of the University of Surrey. He remained in contact with Marian Rejewski until his death in 1978.

6 Celebration at WCC 2018

At the IFIP World Congress in Poznan in Poland IFIP WG 9.7 on the History of Computing held a stream on computing in eastern Europe. One of the most significant events of World War 2 was the breaking of the German Enigma codes. As noted earlier, the contribution of the British codebreakers has been widely described but the work of the Poles has been largely unacknowledged.

The Congress provided an opportunity to put right this omission. The day celebrated the work of three talented and heroic Polish mathematics students from the University of Poznan, Marian Rejewski, Jerzy Rozycki and Henryk Zygalski, who trained in Poznan to become codebreakers and whose work ultimately led to the significant shortening of World War 2. The event attracted significant media interest including TV and radio in both Poland and the UK. The event was also webcast and is currently available online (YouTube [2018](#)).

The one day Bombe stream comprised three lectures and a Bombe Roadshow challenge under the title of “Enigma Live”. The opening talk was by Sir John Dermot Turing who asked the question “Did Alan Turing see an Enigma machine at Bletchley Park?”. The second two talks were by Prof Marek Grajek from Poland. He spoke on the work of the Polish Codebreakers and secondly the proposed Poznan Enigma Centre one of whose main aims will be to promote the interest of young people in cryptography and computing.



Fig. 3. Turing Welchman Bombe used to break the Poznan message

The Bombe Roadshow was a challenge to decode an Enigma message using the Turing Welchman Bombe in the UK. This is a fully authentic replica of the machine originally designed by Alan Turing, enhanced by Gordon Welchman and built by BTM (Fig. 3). It is regularly demonstrated at The National Museum of Computing housed in Block H at Bletchley Park in the UK by the Bombe team of volunteers. The Bombe’s function was to find feasible wheel positions which is a critical and time consuming procedure in finding the key of the day. This process is fully explained in (Turing 2014).

The plan for the event was to send, as an email attachment, an encrypted message with its clear equivalent (or “crib”) followed by another encrypted message whose contents were unknown to the Bombe team. Due to a minor technical fault limiting the Bombe’s operating speed it was necessary to send the crib message ahead of the event. Otherwise the day ran to plan and a successful break was made in the early afternoon when the decrypted message was sent to Poznan from the UK.

References

- Kozaczuk, W.: Enigma. Arms and Armour Press (1984). ISBN 0 85368 640 8
- Perera, T.: Inside Enigma. Radio Society of Great Britain (2010). ISBN 978 1 90508 664 1
- Pidgeon, G.: The Secret Communications War – The Story of MI6 Communication 1939–1945. Arundel Books (2008). ISBN 978 0 95605 152 3
- Turing, D.: Demystifying the Bombe. The History Press (2014). ISBN 978 1 84165 566 6
- Turing, D.: X, Y and Z – The Real Story of How Enigma was Broken. The History Press (2018). ISBN 978 0 75098 782 0
- Turing, D., Grajek, M.: Enigma Live webcast – eight talked including talks. Chaired by Roger G. Johnson. <http://wcc2018.org/Enigma-live>. Accessed 1 Jan 2019

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Professionalism and Frameworks

Moira de Roche^(✉)

Chair IFIP IP3, Johannesburg, South Africa
mderoche@ipthree.org

Abstract. In two sessions the International Professional Practice Partnership (IP3) of IFIP addressed a number of frameworks that provide definitions of ICT competences and typical profiles. These frameworks contribute to establishing an ICT profession that consists of competent and responsible professionals who can demonstrate the necessary skills and competences.

Keywords: Professionalism · Competences · Skills frameworks · Certification · e-CF · SFIA · ACS cyber security framework

1 Professionalism and IP3

1.1 The Importance of ICT Professionalism

Information and communication technologies (ICT) impact almost every facet of personal and business life. Such technologies are key drivers of innovation and of both economic and social progress, making enormous contributions to prosperity and to the creation of a more open world, enabling pluralism, freedom of expression, and allowing people and organisations to share their culture, interests and undertakings worldwide.

Such powerful technologies, and their application, must be driven by competent and reliable professionals who can demonstrate the necessary competences (including knowledge), integrity, responsibility and accountability, and public obligation.

Recognising that ICT is now a global industry, the ICT profession must also be global. It must have clear international standards that accommodate cultural differences in the regulation of professions, which is enhanced by strengthened competence requirements.

1.2 International Professional Practice Partnership – IP3

Through IP3, the International Professional Practice Partnership [1], IFIP established a global partnership that promotes professionalism. By doing so it strengthens the ICT profession and contributes to the development of strong international economies by creating an infrastructure that will:

- encourage and support the development of both ICT practitioners and employer organizations;
- give recognition to those who meet and maintain the required standards for knowledge, experience, competence and integrity; and

- define international standards of professionalism in ICT.

IP3 defines and maintains global standards for ICT and recognises and certifies professionalism. Frameworks underpin the accreditation process, and their use is essential to the maintenance of professional standards at any IT Society or body that is certified.

To carry out its' mission, IP3 works closely with partners who share a commitment to creating a sound global ICT profession. IP3 encourages employing organisations, governments, commercial enterprises and IFIP member societies to join in this partnership through their membership.

IP3 in 2016 launched iDOCED, the IFIP Duty of Care for Everything Digital campaign which promotes trust in ICT, and the duty of care that everyone should have in the digital world. Duty of care goes hand in hand with professionalism, as trustworthiness is an essential element.

2 Professionalism and Frameworks

A number of frameworks has been developed that provide definitions of ICT competences and typical profiles. These initiatives are characterised by an open and inclusive approach, and accredit valuable qualification elements, either national, regional or global. In two workshops at the IFIP World Computer Congress (WCC) 2018 an overview was provided of some frameworks in use, as well as the practical implementations of the frameworks.

2.1 e-CF Overview

[2]: “The European e-Competence Framework (e-CF) provides a reference of 40 competences as applied at the Information and Communication Technology (ICT) workplace, using a common language for competences, skills, knowledge and proficiency levels that can be understood across Europe. The European e-Competence Framework provides a common language to describe the competences including skills and knowledge requirements of ICT professionals, professions and organisations at five proficiency levels, and is designed to meet the needs of individuals, businesses and other organisations in public and private sectors.”

Cleary [3], Deputy Chief Executive of the Irish Computer Society and former chair of the e-CF workshop, explained the state of the ICT profession in Europe, focussing on maturing the profession, with a short-term aim of a fully professionalised sector. The profession must be committed to a relevant body of knowledge, with standardised competences, a commitment to continuous professional development and a clear code of ethics. Progress has been made in achieving these goals.

e-CF is now a European standard (EN 16234-1, April 2016), and work is underway to establish a standardised Body of Knowledge (BoK), education and certification, and a code of professional ethics. There is a standardised set of 30 ICT professional role profiles, which are fully incorporated into the EC ICT Rolling Plan for ICT Standardisation.

2.2 SFIA Overview

[4]: “The Skills Framework for the Information Age (SFIA) describes skills and competencies required by professionals in roles involved in information and communication technologies, digital transformation and software engineering. It provides a framework consisting of professional skills on one axis and seven levels of responsibility on the other. It describes the professional skills at various levels of competence and it describes the levels of responsibility, in terms of generic attributes of autonomy, influence, complexity, knowledge and business skills.”

Seward [5], General Manager of the SFIA Foundation, (Skills Framework for the Information Age), explained the history of SFIA and how the framework is used in the Skills and Competency Management Cycle. A useful description of the structure of the framework was included, which comprises: six categories; 17 subcategories; 102 skills names with associated skills descriptions; and 388 skills level descriptors in the professional skills component. There are seven levels of responsibility, five generic attributes, and 35 attribute level descriptors in the behaviours and knowledge component.

SFIA is developed by industry and business for use by industry and business in the real world. At its heart is experience. A practitioner has a skill or competence because of the experience of practicing the skill in a real-world situation.

2.3 e-CF in an Academic Environment

Bolanowski [6], representing the Faculty of Electrical and Computer Engineering Rzeszow University of Technology and the Polish Information Processing Society IT Competence Council, considered the e-CF in an Academic setting.

An overview was provided of the typical University of Technology graduate, and considerations of the employers’ needs. The university explored “Who is the modern IT Specialist?”. They needed to consider the legal issues ruling university education in Poland. The e-CF was examined considering the point of views of all stakeholders: the student, the university and the employer. Questions to be answered were: What are the possibilities of implementing the e-CF in the university environment in relation to the needs of the job market; How can the students use the e-CF to build and develop their careers; and What are the difficulties associated with the implementation of the e-CF in the university environment?

The technical competences and business/soft competences, required by the job market, were explored. The issue is complicated by the lack of a definition for an IT specialist. Common definitions and terminology are required, and it is hoped that the e-CF will provide at least part of the solution.

From a students’ perspective e-CF helps in organizing the requirements of the job market, it creates a common terminology dictionary, it helps to determine competences and it can help in career planning. For teachers it allows to periodically verify the content of the educational module and to focus not only on technical skills. It may also help to internationalize the education process.

For employers it can improve communication between companies, students and universities, it can help to organize the employment structure. It might also allow a

company to prepare internship programs and to actively participate in the educational process. An additional benefit may be a reduction in the costs of the recruitment process.

2.4 ACS Cyber-Security Framework Overview

This framework, developed by the Australian Computer Society (ACS) as the basis for an extension of the ACS professional certifications scheme and adopted by IP3, was presented by Wong [7], IFIP IP3 Director, and Immediate Past President of ACS.

The following points were covered: Cybersecurity, Privacy and Technological challenges – what are Organisations and Governments seeking; How can we align Business and Organisational priorities with those of security professionals; The Professionalisation of Cybersecurity and Privacy practitioners; Challenges & key issues – is EU GDPR transforming the landscape; What are the repercussions of doing nothing?

The Duty of Care that is requisite for governments was examined, and the rationale illustrated. This includes: Measures relating to the confidentiality, availability and integrity of information that is processed, stored and communicated by electronic or similar means; The implications for government perspectives on cyber law, cyber policy both local and international, how issues and attacks are communicated, offensive cyber security, the cyber economy, cyber intelligence and forensics. Furthermore, there is a global shortage of Cyber-Security professionals, which runs into millions.

Cyber Security, Privacy and Technological challenges include: Definitions of ‘Cyber security’ still unclear; There is a strong demand for Cyber security practitioners but understanding of ‘professionalism’ not explicit; Pseudo Professional Standards proliferate; Cyber security and privacy issues are now mainstream in the boardroom. These challenges together with urging from government, resulted in the development of the Cyber Security frameworks. These are “specialisms” which are in addition to the standard IP3 Technologist (IP3T) and IP3 Professional (IP3P) certifications. In response to this, the frameworks were developed in Australia by the ACS. They are designed to provide a level of Assurance and Trust, and to address the growing shortage of cyber security expertise. The frameworks and related certifications will raise professional standards for cyber security specialists and highlight the Duty of Care for cyber security professionals.

3 Frameworks Implementation

The second workshop included sessions which explored the practical implementation of the frameworks. Ruoff [8], KNVI (Netherlands) explored the usage of the new e-CF profiles and role documents in different settings. She provided an overview of the Professional ICT workforce in the EU, as at 2016. The purpose of the e-CF is to provide a shared language which can be used to address the Skills gap. She provided information about the building blocks of the Framework and went on to explore these in detail. She demonstrated the mapping of the e-CF to SFIA. To tie it all together, she shared several use cases:

- Job profiles for information security 2.0, PvIBQIS
- Supplier Management: KPN consulting IT-CMF –e-CF
- Data Science, EU-Edison project, University of Amsterdam
- E-CF© NEXT, profile tool/assessment of EXIN
- Rake-Shape, blockchain f.e. UWV, LRWA.

Tony Parry, IITPSA, explained how SFA is used for membership grading. He explained that IITPSA uses SFIA as a standardised approach, that is consistent, fair and aligns to IP3 and the South African Qualifications Authority requirements. It is used for: the Professional Designation (PMIITPSA) which is IP3 accredited and SAQA registered; peer reviews; Critical Skills Assessments (foreign worker work permit requirement). The philosophy is “Assessing each case on its merits”.

Dębski [9], PTI-IT Competence Council, Ministry of Digital Affairs examined how the e-CF can be used in the education system for ICT Professionals. The e-CF should be considered in Computer Science Education for high school and vocational school students, as the foundation of the creation of the future ICT Professionals, as well as their development. Having discussed the skills requirements with all stakeholders, they were very pleased to discover the e-CF. Using the e-CF and its 40 competences helped to speed up the work of developing curricula. They took the competence framework, went through its 40 competences and on this base we actually created a common language. The power of the e-CF is that it relates to real life and real labor market.

Wong [10] provided insights of the Cyber Security Framework in Action. He examined the situation around the world, especially the effects of GDPR. The biggest threat is the severe shortage of skills in the Cyber Security space globally, citing examples showing the estimates of 1.8 million people. The EU is leading the world in legislating to protect and provide access to personal data with its EU General Data Protection Regulation (GDPR), which replaced the 1995 Data Protection Directive in May 25, 2018. The implications are far-reaching, affecting with an establishment in the EU or that offer goods and services to business or citizens of the EU, or that monitor the behaviour of individuals in the EU may need to comply. There are significant penalties for non-compliance with fines up to €2million, or 4% of global turnover. GDPR specifies job designations related to compliance officers, and the requisite skills for those in these roles. The ACS developed the Cyber Security Framework:

- Designed to provide a level of Assurance, Trust and to address the growing shortage of cyber security expertise.
- Launched by Australian Minister Assisting the Prime Minister for Cyber Security, the Hon Dan Tehan in Canberra in Sept 2017.
- Adopted by IFIP IP3 as a new specialism certification for member societies around the world.
- ACS support for the implementation of the Australian International Cyber Engagement Strategy announced by the Hon Julie Bishop MP, Minister for Foreign Affairs in October 2017.
- Raise professional standards for cyber security specialists.
- Highlight the Duty of Care for cyber security professionals.

Reflecting the multi-disciplinary nature of Cyber Security, flexibility is built into the certification for Technologists (SFIA Level 3) and Professionals (SFIA Level 5). Professionals who have achieved the Cyber Security Professional certification come from the aviation, banking and finance, audit and risk, consulting, and healthcare industries.

Adrian Schofield, Chair IP3 Standards and Accreditation Committee, explained how frameworks are used to ensure the trust aspect of accreditation (video presentation). Framework ensure that levels are benchmarked – irrespective of the framework used the skills levels and competencies are at a similar level.

4 Follow Up

At the end of the workshop, the speakers and audience considered how it can be ensured that all frameworks are mapped to each other, and the work that needs to be done to realise this goal. We accept that more than one framework is being utilised but encourage new entrants to use something that already exists, rather than create a new framework. Having too many frameworks causes confusion and it duplicates work. Mapping and customisation are better options.

An IP3 task force has been created to develop a project plan to carry out this work in collaboration with all key players. It is hoped that this project will be funded. IP3 call for all interested parties to join us in this work to ensure that the process is inclusive and representative. Contact mderoche@ipthree.org for more.

References

1. IFIP IP3. <https://www.ipthree.org/>
2. <http://www.ecompetences.eu/>. Accessed 18 Jan 2019
3. Cleary, M.: <https://www.ipthree.org/wp-content/uploads/Mary-Cleary-e-CF-and-TC-428.pdf>
4. <https://www.sfia-online.org/en>. Accessed 18 Jan 2019
5. Seward, I.: <https://www.ipthree.org/wp-content/uploads/SFIA-Overview-Ian-Seward.pdf>
6. Bolanowski, M.: <https://www.ipthree.org/wp-content/uploads/MB-Frameworks-in-an-Academic-setting.pdf>
7. Wong, A.: <https://www.ipthree.org/wp-content/uploads/Cyber-security-Framework-overview-Anthony-Wong.pdf>
8. Ruoff, L.: <https://www.ipthree.org/wp-content/uploads/eCF-Implementation-Liesbeth-Ruoff.pdf>
9. Dębski, B.: <https://www.ipthree.org/wp-content/uploads/e-CF-Education-System-B.Dębski.pdf>
10. Wong, A.: <https://www.ipthree.org/wp-content/uploads/Cyber-Security-specialism-framework-in-action-Anthony-Wong.pdf>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Should Artificial Intelligence Be More Regulated?

Panel Discussion

Leon Strous^(✉)

De Nederlandsche Bank, Westeinde 1, 1017 ZN Amsterdam, The Netherlands
strous@iae.nl

Abstract. Artificial Intelligence (AI) can and does bring immense benefits in all sorts of areas. But it also introduces (new) risks. Is more regulation needed? In order to answer this question arguments pro and con were presented by four panel members and discussed and challenged by the audience. Many issues were raised, ethical principles, the obstacles that make it hard to draft good legislation. We don't want to stifle innovation or deny society the benefits of these technologies by excessive regulation. A distinction is made between science (research) and the application of AI technologies. Comparisons with other sectors and technologies are made to see whether parallels can be drawn.

Keywords: Artificial Intelligence · AI · Regulation · Ethics · Ethical principles · Liability · AI science · AI applications

1 Introduction

Many discussions are taking place at the moment about Artificial Intelligence (AI), about ways AI may benefit mankind and about risks of AI. Autonomous cars, automated trust assignment to individuals, and autonomous weapons are only a few examples how AI can change our life. Some people warn us that AI can be even more dangerous than nuclear power. On the other hand it seems impossible and undesirable to stop development of AI and its applications. Thus, the question arises what should be the role of governments. Should AI be more regulated with respect to research and/or its usage? This question was addressed at WCC 2018 in a panel discussion with four panel members:

Ulrich Furbach, University Koblenz-Landau, Germany,
Eunika Mercier-Laurent, Lyon III University, France,
Chris Rees, British Computer Society, UK and
Jerzy Stefanowski, Poznan University of Technology, Poland.

An audience of 40 participants actively engaged in the discussion. The session was recorded on video [1].

2 Arguments Pro and Con

To start the debate, two panel members presented arguments in favor of more regulation and two members presented arguments against more regulation.

Arguments in Favour of (more) Regulation

AI can and does bring immense benefits in all sorts of areas: in cancer diagnosis, mind diseases, caring of elderly, and many more will follow such as autonomous vehicles. We don't want to stifle innovation or deny society the benefits of these technologies by excessive regulation. Any regulation should be risk based. If the risk is low, regulation should be avoided. If it is high, the application should be regulated. That's how it is in the non-AI world and the AI world should be no different. Further we have to recognize that drafting regulations for new and fast developing technologies such as AI is difficult. There is a risk of building assumptions and language in the regulation that don't stand the test of time.

Our starting point is ethics. The implementation of AI systems, including of AI driven robotics, poses a number of ethical challenges. A non-exhaustive list includes:

- reliability and safety of complex systems,
- bias in systems and bias in the data,
- black box systems that cannot explain or justify their decisions,
- the allocation of responsibility for failure,
- malicious use of AI and lethal autonomous weapon systems,
- the destruction of jobs by AI,
- the protection of privacy.

The question is: where can we rely on the ethical actions of developers and users of AI and where is this clearly not adequate and therefore regulation is necessary. Some of the applications of AI are in domains that are already regulated and have a long history of regulation, medicine and finance are two obvious examples. But the existing regulations may not cover the application of AI. These regulations may need to be enhanced to prevent harm to patients or unfair financial practices. Autonomous cars should not be allowed to go on the road until there is an agreed allocation of responsibility and therefore liability for harm. You need a third party insurance that covers the driver. But in an autonomous car there is no driver. So who is responsible/liable: the manufacturer of the car, the manufacturer of failing components, the car salesperson, the owner of the car? And what if the car is hacked? Or if software updates have not been installed? There is no doubt that existing regulations do not cover autonomous cars and that this is needed.

Many people have already asked for a ban on autonomous lethal weapons, comparable to nuclear and chemical and biological weapons. While controlling adherence to such a ban may be difficult and sometimes such weapons are used nevertheless, laws and regulations have a powerful effect on the public opinion.

Economists predict a growth in jobs due to AI, but only in the long term. In the short term jobs will be lost. Regulation may be needed to provide funding for retraining employees for a new job.

AI can be used to de-anonymize personal data that has been anonymized. GDPR may already be a good step in the right direction restoring the control over personal data to the owner instead of the company. However the protection in the GDPR against AI based use of personal data is weak and needs strengthening.

Machine learning and AI systems are complex systems. In a number of application areas we should think about regulation. Consider machine learning and AI systems as a product and therefore regulation is focused on the application of AI, on the product. In the medical domain we see systems that can make prognoses and by doing that impact people's health. It is important that the system can not make errors. It is the task of the producer/vendor to take care of that. Compare it with the process to get a new drug (medicine) accepted. Strict procedures and tests take place before the new drug is allowed to be put on the market. Producers of AI should provide assurance that their product is working correctly and this should be enforced through regulation. Regulation does not always have to be laws, it also can be community agreed rules and processes or evaluation and certification. Another element concerns the question whether an AI system should be able to explain the decision it took. For some domains and applications this may not be necessary, for others it is, think about legal decisions (e.g. AI supported court cases). It should be mandatory for such systems to be able to explain. That may not be easy, when is an explanation clear enough, what is the context.

Another issue is intellectual property rights. Advanced systems can write poems and stories or compose music. Who owns this and benefits from the profits this might generate? Regulation may be needed to clarify such rights.

Arguments Against (more) Regulation

Regulation of AI may be undesirable and extremely difficult if not impossible. A number of questions support that position:

- Regulation may work in a normal ethical society but how can we regulate a society that is composed of robots and humans.
- If regulation is used to prevent machines from doing "something foolish", who decides what is foolish?
- We live in a business driven world. What will happen if we try to regulate the market giants? They will move to countries without regulation.
- If you look at military use of AI, that is big business with powerful people behind it. Extremely difficult to regulate.
- How will regulation be effective in data protection if people are willing to provide their data voluntarily to companies.
- What about regulation and the creativity of the researcher. Efforts to regulate this without proper understanding how researchers work may lead to disasters.
- How to regulate a robot from learning. How to say to a robot what he can and what he should not learn.

If regulation is needed, it could be considered to not only look at legal regulations but also at initiatives about principles. An example of these are the Asilomar AI Principles. Industry giants and experts such as Elon Musk and Stephen Hawking have advocated for humane and "safe" robotics. Along with hundreds of researchers and

experts in the fields, they have proposed 23 “guiding principles” that will ensure the development of AI for the benefit of mankind [2, 3].

Although one might in principle be pro regulation of certain aspects of AI system, it simply seems to be impossible. The problem is we don’t know what it is, an AI system. AI is not a monolithic system, it is in other systems, it is in our cars, in our search engines, in our shopping cart. AI is a functionality of existing systems. It is completely impossible to control the development of these techniques and also it is impossible to control its use. Other areas that have been regulated also show this. Two examples from the weapons industry. Nuclear weapons, we all know how difficult it is in certain parts of the world to control the development of nuclear weapons. It is a highly political issue. The other example is chemical weapons. They are banned and nevertheless used. It is impossible to control the use of technology and it is impossible to control the development of technology. We also should not want this because technology is a driving force of our society. We want to learn more. We shouldn’t stop science or regulate science. There may be some exceptions with respect to ethical issues. We don’t know exactly what to regulate and how to regulate it. The United Nations does not succeed in getting a letter signed by all countries about the goal of a ban on lethal autonomous weapons. Some countries have major interests in such an industry or other arguments for not signing it.

Another example concerns autonomous cars. Where liability and insurance issues might be regulated there are also ethical issues. The German government drafted a report saying that an autonomous car should never be able to face an ethical dilemma situation. That is impossible, it would be similar to saying that human drivers should never face an ethical dilemma. Furthermore the report argues that algorithms should be checked and that self-adapting systems should not be applied in autonomous cars. That is also strange, an autonomous car should learn by driving and adapt. It is also unimaginable that a human driver would be forbidden to learn from mistakes.

In a distant past when cars were just introduced, there was a rule in the UK that in front of a car there should be someone walking with a red flag to warn people. Perhaps we should use a red flag to warn (or better: make people aware) that you are dealing with an AI system. A Blade Runner situation where it is difficult to distinguish humans from machines should be avoided. Regulation might be helpful for that.

3 Summary of the Debate

During the debate the arguments pro and con were both challenged and supported and some new issues were raised. This chapter provides a selection of the main topics discussed. Sometimes in a Q&A format, sometimes just as additional remarks.

AI tools, systems, technology could/should be regulated as human beings are also regulated. What about AI as a scientific discipline, should that be regulated as well? In a sense this is already regulated in the same way as other scientific disciplines like medicine and genetics. When applying for research funding for instance the request has to be judged on many aspects including ethical issues. What should not be regulated are the goals to pursue with research.

A comparison with the regulation of the Internet can be made. We now realize that we were too late thinking about regulation of the Internet when the Internet was created and that makes it difficult to repair it now. Maybe it is also due to the way scientists think. The benefits prevail, especially in areas where AI can assist professionals such as medical doctors who are already overburdened to take over part of the routine work. And focusing on the benefits, the risk of abuse of technology that is created with the best of intentions might be overlooked.

Regulations should be in place in certain areas but an additional question is who will be responsible for those regulations. If it is the lawmakers do they know enough about the topic to draft good regulations. A lot of poor law is written because of insufficient knowledge of the subject matter. Society/lawmakers lag behind with respect to technological developments. We as professionals at the forefront of these developments are better placed to judge where these developments might lead and what might be an appropriate societal safeguard long before the lawmakers can make those decisions. This means that we as an IT community have an obligation to engage with legislators to support them in drafting decent legislation. We should at least make an effort to be involved.

The issue was raised that a request for funding of scientific research usually has to pass via ethical committees, because funds are tax payers money. That is not the case for research and product development done by industry. AI as technology may perhaps not face ethical questions but the applications do. Is the current ethical oversight (for academic funding) sufficient? During many years of AI research ethical questions never popped up, we researched nice technology. Because there were no real-life applications. Now this is changing for instance with autonomous cars. And that introduces also the question of impact of the application. The impact of an application is often not or rarely assessed before selling or using it.

An interesting perspective was mentioned from a small and medium sized enterprise point of view. When you develop a new product, regulation in the beginning is an obstacle and difficult. However, it can also mean a benefit if you can advertise that your product meets certain regulations. And the competition has to keep up with that. Also good for consumers who can see that a product meets certain requirements as laid down in regulations.

Regulation in the globalized world is difficult. It is not enough to have regulations in a country or a region. However, so far regulation on a global scale, for instance via UN, is not successful. If we want to regulate a borderless development such as AI, we need to do that on a global scale otherwise it is meaningless. The statement that regulation will only work if it is on a global scale was challenged. Take for example the argument that companies will go to a country that is not regulated. That will not help them because they can maybe produce the product in such a country but if they want to sell it in a country that has regulation, it will not be able to do so unless it complies with the regulation. GDPR is a good example. US companies have to comply with GDPR if they want to do business in the EU. It can work well even if it is jurisdiction based.

Another talk at WCC about shifting identities triggered an issue of importance to AI. Identities touch a multidisciplinary field. Regulating this part of science also means that we need to be very clear where we want to go, what we want to be in the future. It was mentioned that perhaps a link can be made to the work on consciousness.

Psychologists and philosophers are trying to find out what does it mean for humans to have consciousness. As developers of AI we are working on AI systems that have a kind of consciousness. A German philosopher Metzinger argues that we should never try to bring consciousness into an artificial system because then we would be able to bring harm to them and we are not allowed to do harm to other human beings.

It is our duty as IT professionals to explain AI to people. We should be able to understand how algorithms work. And to explain the choices that have been made in designing the algorithm and the effects the algorithm may have.

4 Conclusions and Follow up

While there was not a complete agreement on everything and the outcome of the discussion was not fully conclusive, in general a broad consensus could be noted on a number of issues.

Artificial Intelligence is a broad term that includes science, technology, applications and products. AI can bring benefits but it also can introduce (new) risks. The answer to the question “should it be (more) regulated” depends on a variety of aspects, it can’t be a yes or no. Consider things case by case. The term regulation is not precise, it can mean a law but it can also mean mutually agreed rules and procedures.

We do not have clear easy answers but we should make efforts and increase awareness. We should debate and work on documents to indicate critical points. We cannot control everything. You should know who you are talking to. Difficult, challenging but not a reason for not trying.

It is important as professionals to engage in discussions like this. We as an IT community have an obligation to engage with legislators to support them in drafting decent legislation.

We should develop methodologies to certify AI products. There is a role for IFIP and other professional societies to think about ways about how to define workflows for approving AI based products.

April 25th 2018 the European Commission issued a Communication COM (218) on Artificial Intelligence for Europe [4]. This Communication sets out a European initiative on AI, which aims to ensure an appropriate ethical and legal framework. The Commission will (selective quotes):

- *set a framework for stakeholders and experts to develop draft AI ethics guidelines, with due regard to fundamental rights;*
- *issue a guidance document on the interpretation of the Product Liability Directive in light of technological developments. This will seek to ensure legal clarity for consumers and producers in case of defective products;*
- *publish a report on the broader implications for, potential gaps in and orientations for, the liability and safety frameworks for AI, Internet of Things and robotics;*
- *support research in the development of explainable AI and implement a pilot project proposed by the European Parliament on Algorithmic Awareness Building, to gather a solid evidence-base and support the design of policy responses to the challenges brought by automated decision-making, including biases and discrimination.*

This Communication addresses precisely the issues raised in the discussion. It also invites stakeholders to participate in the efforts. Let's contribute to these and other efforts in the world. There is a need and an opportunity for us as IT professionals to pick up the challenge and to continue the discussion. There is momentum now, let's not waste the opportunity. We don't want to observe in ten years' time that we again missed the boat (after the Internet). We also have to research some fundamental questions, where do we want to go with regulation, where do we want to go with applications, who do we want to be. This is an appeal to all participants and readers who are interested in continuing this debate in a search for guidance. If you want to get involved, let me know. Send an e-mail to the address at the start of the paper.

References

1. WCC 2018 session should AI be more regulated? Video recording. http://www.wcc2018.org/movs/oxford_debate_rf27.mp4. Accessed 21 Jan 2019
2. 23 AI principles, news article. <https://www.natureworldnews.com/articles/35633/20170215/23-principles-ai-stephen-hawking-elon-musk-experts-pitch-rules.htm>. Accessed 21 Jan 2019
3. 23 AI principles, Future of Life. <https://futureoflife.org/ai-principles/>. Accessed 21 Jan 2019
4. Artificial Intelligence for Europe, EC Communication COM (2018) 237, 25 April 2018. <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>. Accessed 21 Jan 2019

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



**IFIPIoT 2018 Invited Papers:
Keynote, Panel Discussion**



The Internet of Things is Hardly About Technology

Kees van der Klauw^{1,2}✉

¹ Alliance for Internet of Things Innovation, Brussels, Belgium
chair@aioti.eu

² InnoAdds, High Tech Campus 27, Eindhoven, Netherlands
kees.van.der.klauw@innoadds.com
<http://www.aioti.eu>, <http://www.innoadds.com>

Abstract. Most definitions of the Internet of Things (IoT) take a technology perspective, referring to connected devices exchanging data with each other and with higher levels, establishing autonomously operating systems. From a business perspective, IoT can be regarded as a business transformation, driving commoditization or even threatening conventional businesses, providing opportunities for product and process improvements and opening perspectives for services business or entirely new businesses based on data acquired by IoT. While the development of the Internet of Things resembles in many aspects the characteristics of for example semiconductor and IT platforms, there are some essential differences, mainly in scale and scope, that make the kickstart of a successful IoT more complex. This paper addresses key elements of a successful deployment of IoT and also key roles for IT professionals in IoT.

Keywords: Internet of Things · IoT · Platforms · Open innovation · Ecosystems · Digitization

1 Introduction

Most definitions of the Internet of Things (IoT) take a technology perspective, referring to connected devices exchanging data with each other and with higher levels, establishing autonomously operating systems. Such definition easily fits in with other technological developments that are driving the digitisation of our world. Sensing, ubiquitous communication networks, information systems, data analytics, artificial intelligence, robotics, edge and cloud computing are all linked to the development of IoT, either as an enabler or leveraging IoT for new applications.

From a business perspective, IoT can be regarded as a business transformation, driving commoditization or even threatening conventional businesses, providing opportunities for product and process improvements and opening perspectives for services business or entirely new businesses based on data acquired by IoT. New economic powers and regions will emerge, others will lose relevance.

There is no doubt that IoT will be a driver of socio-economic change like the industrial revolution was. Our society will increasingly be run by autonomous systems, self-learning robots will assist people not only in manufacturing, but also in their daily

life. Resource management (energy, food, water) will be based on IoT, combining data from numerous sensors and actuators and systems in the field, providing real time insight and steering optimized flows. Curative healthcare will be replaced by a continuum with preventive monitoring, highly automated treatments and home care. Systems of systems will autonomously run operations in energy, mobility, cities, buildings, industry...

Jobs will disappear while other jobs are created. Education will change to life-long learning while new creative fulfilment of free time will be required. And since the speed of change is fast, people will feel uncertain, distrusting new technologies.

Don't worry about the definition of IoT. IoT is all of it and it is about time that we take an integral, holistic approach while 'separating concerns' to make it work for all.

2 A Short History of the Future: Enabling Platforms

Already in the 80-ies have we seen the need for such 'separation of concerns' approach in the electronics industry, although predominantly in the technology area. Socio economic considerations were lagging, as usual.

In the 80-ies the electronics industry was predominantly analog and the design, manufacturing & application required (scarce) deep and wide knowledge. Chip application engineers, designers and technologists were very much from the same origin and closely working together.

But once IC technology was mastered, one could massively apply transistors as switches and a digital technology platform was established. While one breed of engineers worried about making a good transistor switch and replicating them hundred thousand times without flaw, a new generation of engineers that never had seen a transistor was enabled with tools to create complex computing circuits with those large numbers of transistors, not requiring deep transistor and technology knowledge but using instead simulation models ('digital twins'). Separation of concerns.

This created a whole new framework of knowledge and associated professions, but it also bifurcated IC business in fab and fabless companies, all enabled by digital technology platforms.

This bifurcation has repeated several times, all driven by the same principle of establishing a platform that enabled the development of new knowledge, professions and business on it, separating concerns from the layers below.

Integrated Circuit Designers started using standard building blocks and those building blocks were combined in standard IC's such as, microprocessors, microcontrollers and communication chips. Again establishing a digital platform enabling a new breed of engineers to run away with them building programmable computers, communication networks and automation and control systems.

Those systems needed standard SW operating systems and standard application software packages in order to have them deployed massively and a new platform layer, this time in SW was established, based on the same economy of scale as Moore's Law driving the bottom of the pyramid. Large software stacks have high creation and maintenance cost and therefore application toolsets enabled a new breed of SW

application engineers to deploy and configure those standard software packages in many domains. This role is largely filled in by many small and medium enterprises.

While this development has disrupted businesses, it has created enormous economic value overall. It has been a strong driver of ‘Moore’s Law’ in the sense that it created the economic justification for the technological advancements in IC technology and on higher levels in SW. Without platformisation, the market would not have developed and Moore’s Law and SW growth would have been stagnating because of lacking returns.

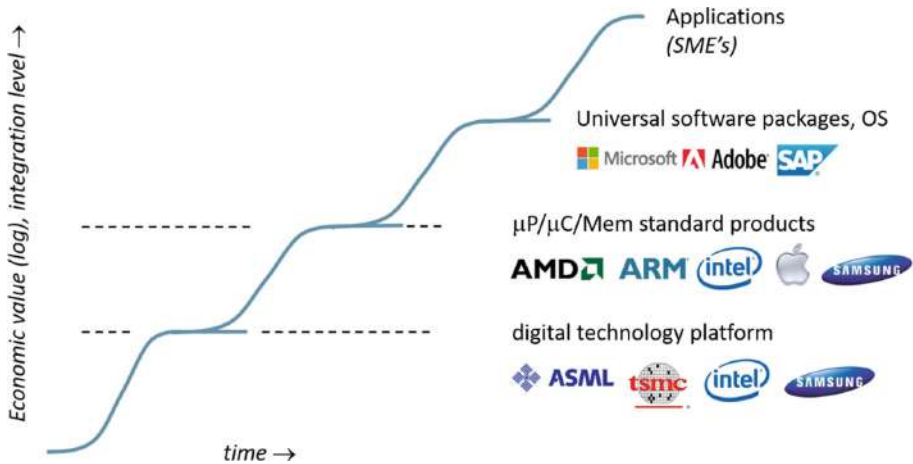


Fig. 1. Schematic representation of technology platforms developed over time in the IC/ICT domain enabling higher integration layers and economic value. Some key players on various levels are illustrated (non-exhaustive)

Figure 1 expresses this stacking of platforms over time, while the overall value exponentially grows with each layer. This is also reflected in the number of jobs that is anticipated in the software industry vs. the hardware industry, being orders of magnitude larger. Particularly in the SW application industry have many smaller companies developed that are configuring and customizing implementations of standard software packages such as ERP systems and databases. This being the case, the economic value is still enabled by the hardware industry.

The rapid development of digital platforms has posed a dilemma for many companies active as they had to decide on their position in this developing value chain. IC manufacturers had to consider focusing on technology or design (separating concerns), or both. Generic SW companies (ERP, OS, DB) had to decide whether to support all its implementations or focus on one of the two. And life is very different above and below:

- Very different perspectives;
- Very different lifecycles;
- Very different competencies required;
- Very different economics (business control points).

But the choice that companies had was enabled by a ‘common interface’ in terms of language, protocols, simulation models... ‘digital twins’, creating a strong interdependency (Fig. 2).

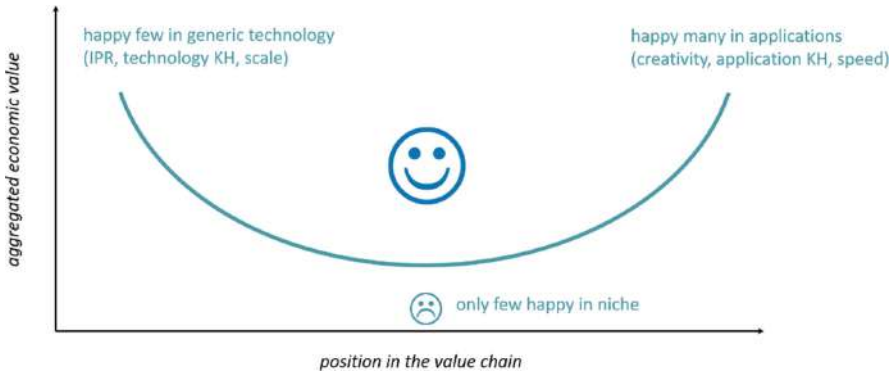


Fig. 2. The smile curve, picking your place in the value chain. In general the economic model on the left side is fueled by high volume, high investments and on the right side by lower volume, high diversity. The number of players on the left side is substantially lower than on the right side, leaving room only for niche players with customized technology in the center.

Already in the 90’s was the so-called ‘smile curve’ introduced to express the dilemma of positioning oneself in the value chain. The vertical axis reflects the economic value that a company may create and the horizontal axis the (upward) movement in the value chain. As stated earlier, the total value of all companies is substantially larger on the right hand side but this value is created by many more application companies. There is only room for fewer players on the left hand side where the rules of the game are in general scale and high investments. The center reflects that you can’t be both and there is only room for niche players with relatively low impact. That does not imply that representatives of the same (vertically integrated) conglomerates can show up on the left and the right... They can, but within those companies they will be very different animals requiring dedicated management and business controls.

3 The Internet of Things: Some Things are Repeats of History

Like the stacked platform picture that one can draw for the IT technology and applications ecosystem, one can create such picture the Internet of Things (Fig. 3). Clearly the IC technology stack is deeply embedded in the first layer of this IoT ecosystem, creating intelligent devices that are able to communicate. They represent (embedded) systems on their own.

Figure 3 does not aim to depict a standard way of looking at the IoT but one could recognize a few layers of the OSI model in it. The main thing is that it represents a very

strong way of ‘separating concerns’ as the various layers have so much specialism in technology, business models, competencies and culture that one could regard it as different worlds that somehow are intimately dependent on each other.

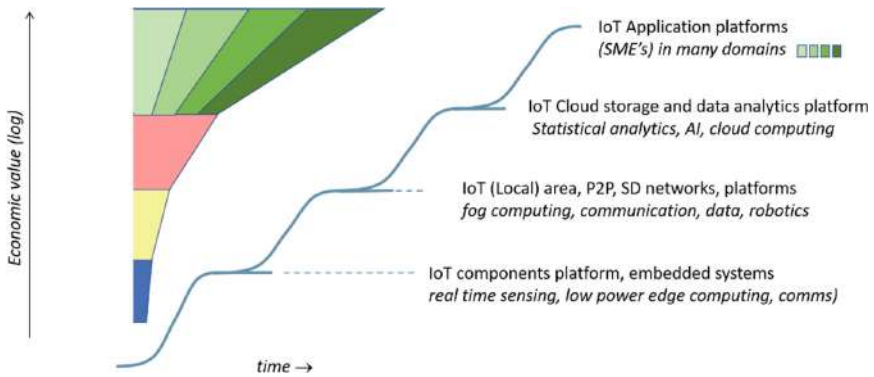


Fig. 3. Schematic representation of some key platform layers in the development of Internet of Things. The economic justification of enabling (lower) layers comes from many applications in various domains, each requiring expert knowledge

The picture also represents different architectural choices when it comes to data processing, that can be executed in end-nodes of the IoT (edge computing), in intermediate nodes, often representing local area networks and servers (fog computing) and central information processing often in the cloud. Hybrid forms are the most likely candidates for a solid implementation of IoT as the non-functional characteristics of centralized and decentralized systems are very different. We will go a bit deeper in this in the next sections.

The higher one gets in this picture, and the higher up in the value chain, the more diversification is taking place and knowledge of the application domain becomes essential. While this was already the case in IT for e.g. bookkeeping, for CAD, for control systems, it is taken to an extreme in IoT. IoT will be much more pervasive in society and one cannot create meaningful IoT applications without a solid understanding of the socio economic impact of it. This is where technology push ends.

Only by enabling an ecosystem that is deeply involved in applications in real life can a scale be created to justify the investments in the underlying layers such as communication networks and software systems. The key question is whether we are on track to create such ecosystem.

4 The Internet of Things: Some Things are Different

While the development of the Internet of Things resembles in many aspects the characteristics of the semiconductor and IT platforms as discussed in Sect. 2, there are some essential differences, mainly in scale and scope, that make the kickstart of a successful IoT more complex.

First of all has the investment scale gone up, for example in communication infrastructures. While the investment in 2G communication networks could be justified by just mobile phone traffic and 3G networks by increased data use on smart phones, future networks will be based on serving trillions of IoT devices in many different application domains with different requirements that all have to be developed.

But the same holds for platform developments in e.g. smart mobility or smart cities. They will require large investments that can only be earned back with many different application use cases. There is no single use case anymore that justifies the investment. This brings a higher level of uncertainty, that has to be countered by a strongly orchestrated approach for which we see 3 variants in today's world:

- Orchestration by governments;
- Orchestration by powerful (vertically integrated) companies;
- By collaborative platforms and ecosystems.

Secondly, IoT technology requirements differ from general Internet requirements that all require dedicated developments. Some of the most obvious ones are:

- Very low standby and communication power consumption in devices;
- Flexible bandwidth allocation (from very low to high);
- Very low cost per node serving extremely large numbers of nodes;
- Flexible and programmable communication layers (P2P, local, central);
- Extremely low latency for several applications;
- Extremely high reliability and resilience for several applications;
- Strong security and privacy requirements;
- ... several other non-functional elements such as sustainability.

Serving some of these requirements will require fundamental changes in the architecture of next generation internet.

Thirdly, the higher one gets in the value chain, the wider the scope that needs to be orchestrated/managed and deep involvement of application domain specialist beyond technology will be crucial for IoT, more than it was for ICT development:

- A very wide range of application domains such as Smart Cities, Homes/Living, Healthcare, Farming, Energy, Mobility, Water Management, Industry... and more to come;
- Many technologies in HW, SW, Communications, Systems Engineering, Robotics, Sensing, Data Analytics, Hypercomputing, User Experience... and all non-functional requirements such as security and privacy;
- Legal and liability aspects;
- Socio-economic and human aspects;
- Education and training;
- Use case development.

The overall orchestration in scope, complexity and uncertainty of all these elements and the establishment of platforms that allow a 'separation of concerns' is probably the largest challenge for the development Internet of Things in Europe, more than the individual developments in technologies.

5 Key Elements for a Successful Deployment of IoT

The key issue with IoT is not technology, but the fact that

*applications and platforms are insufficiently established. . .
because supporting platforms are insufficiently established. . .
because their justification by applications is insufficiently established*

This is depicted in Figs. 4 and 5. Figure 4 represents the idealized stack of ‘smile curves’ of players in the components and IoT devices domain, the networks domain, the data domain and the application domain. As stated before, this is not necessarily the only way of breaking down the IoT value chain, but it illustrates that matter. The value is increasing exponentially along the chain and within each domain, one has to pick its position, require a high degree of focus and specialization.

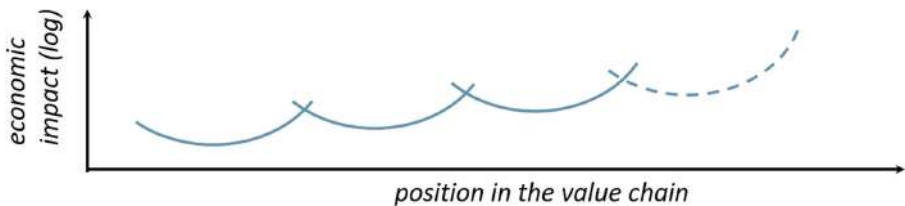


Fig. 4. The idealized stack of smile curves for an IoT value chain. Common interfaces and well connected platforms establish a high economic impact

This ‘ideal’ situation is emerging in ecosystems driven by government in China or by very strong platform companies in US. Note that the left hand part of the value chain does not require strong orchestration as industry platforms have very much established itself, but the challenge starts already at next generation communication networks.

However, a more realistic picture today which holds for Europe is depicted in Fig. 5. Not being driving by a central government powers or by large platform players, there is a lack of established data and application platforms, orchestrated and managed standard interfaces and in essence a lack of common goals. Even though individual expertise and goals of companies may differ, they require a common goal in IoT platform developments to be successful for their individual successes. Without such platforms, their individual value creation will become obstructed.

Not having a central governmental power for orchestration, nor extremely powerful platform companies established in Europe, the question to companies in this ecosystem is therefore very much: *Do you bet on your monopolistic do-it-yourself power or on your collaborative power for IoT?*

The only alternative way forward in establishing the required ecosystem seems to be the creation of *collaboration platforms*, uniting individual players around the common interest. The establishment of such platforms in Europe first of all assumes a sense of urgency in European industry that no single company/organisation can manage the whole stack on its own related to:

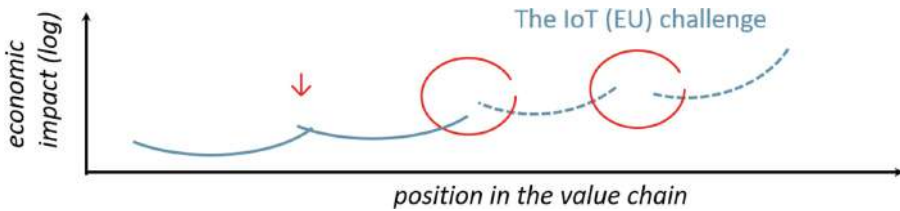


Fig. 5. A representation of today's ecosystem in EU Internet of Things, with data and application platforms and their orchestrated interfaces lacking, making the application value creation speculative and also reducing the economic value of lower layers

- Investments required and need to focus for economy of scale
- Knowledge and competency scope required for different parts
- Culture required for different parts

and a willingness to work together in an open innovation ecosystem that:

- Separates concerns ('mind your own business') building on strengths;
- Shares the common elements that don't differentiate the one company from the other (protocols, standard functions, infrastructures, architectures);
- Creates a large degree of interoperability.

Such ecosystem, although much more difficult to establish than alternatives that work under singular top-down control, can even prove to be more stable and attractive in the longer term because of:

- Involvement of a wider group of stakeholders, leading to higher acceptance and trust by society and governments;
- Better attention to non-functional aspects of IoT, in particular security, privacy, portability, resilience, flexibility;
- Faster development of use-cases, particularly when SME's and creatives gain access to these platforms. Their applications can greatly contribute to the justification of platforms whereas the creation and maintenance of platforms generally requires the skills and scale of larger companies.

And so, the key issue with IoT development in Europe is not technology, but a lack of:

- Collaboration between industrial players;
- Collaboration across functional silos;
- Understanding that economic justification of IoT infrastructures comes from multiple use cases in multiple domains, requiring an integrated approach.

It is not just the creation of platforms that needs a justification from many use cases in many domains, but very much the maintenance of such platforms, requiring a 24/7/365 high performance in the operational management of cities, homes, healthcare, energy, mobility etc.

Since the economic justification for any platform comes from *collectives* of applications, another key element in the deployment of IoT is the selection of those

applications or use cases. This comes with a great deal of uncertainty that can only be managed by applying large scale experimentation in real scale environments. The mindset for the linked development of platforms and applications should therefore be:

*Many applications will fail, get used to it!
Several applications will be successful, count on it!*

Unlike technology development that usually takes place in laboratories, the development of uses cases takes needs to take place in cities, homes, healthcare ecosystems etc., involving:

- A real world environment
- Early involvement of end-users and key stakeholders beyond technology
- Expectation management (including non-functional aspects)
- The notion that early failure on any aspect is valuable learning

Governments have an important role to play in facilitating such real-world experimentation, e.g. by using innovative procurement procedures instead of traditional buying of upfront specified solutions against lowest cost.

Particularly the involvement of societal stakeholders in this experimentation is crucial in order to get valuable feedback on implementations but also to gain trust and address critical concerns in society regarding the impact of IoT applications and new technology that could hamper a successful deployment of IoT, even though technology works perfectly. In that sense should experimentation include e.g. legal and social aspects.

Lack of trust is considered the largest inhibitor of IoT deployment and lack of trust is strongly related to the non-functional aspects of IoT such as privacy and security. Trust is perception that cannot always be addressed by technology. Education and involvement of people in the development of IoT is essential and new knowledge and insights will be developed that in the end could pay off for Europe running a human centric socially embedded IoT.

The Internet of Things development is largely a self-fulfilling prophecy creating higher value for all, provided we:

- Create a collaborative platform in many aspects beyond technology
- Involve a wide range of (societal) stakeholder in an early stage
- Are prepared to experiment, fail and learn
- Are prepared to enable higher value creation before re-distributing it.

6 Key Roles for IT Professionals in IoT

From the sections above, one may anticipate also large changes in certain professions, particularly the ones that operate on interfaces of different disciplines or across different domains and in applications. Other, more specialised professions will not fundamentally change even though underlying technologies will evolve.

Professionals at the end of the chain, in applications, will be increasingly confronted with the enormous opportunities of IoT. This is already happening in e.g. Smart Farming, where the new generation of farmers is strongly involved with the latest technologies in sensing, data analytics and automated growing control in e.g. city farming. In many other domains will the traditional craftsmen become users of intelligent systems, robotics and data that will on the one hand replace many of their traditional work and on the other hand enable them to do ground breaking new things. This is a large part of the social transformation.

Many more IT professionals will be required that operate on interfaces, bringing a wide scope of knowledge and experience or able to link with application domain specialist, talking ‘farmers language’ with IoT farmers, ‘transporters language’ with IoT transporters etc. Several universities and colleges have already recognised this need and started educational programs for:

- T-profile Engineers, ‘platformers’ integrating many functional and non-functional aspects of platforms with strong IT components (communications, data, embedded systems). Typically these professionals need strong architecting and platform management skills;
- Π -profile Engineers, linking with other disciplines (legal, social) and/or engaging with application domains, requiring non-technical skills and application knowledge.

These profiles are schematically depicted in Fig. 6. I-profiles these days form the majority of ICT profiles, focusing on dedicated subjects and specialism in the ICT landscape. But there is a limitation to what a single mind can do in this increasingly complex world of systems. It is unlikely that ‘I can do it all’ and increasingly we need T- and Π -profile Engineers integrating vertically and horizontally.

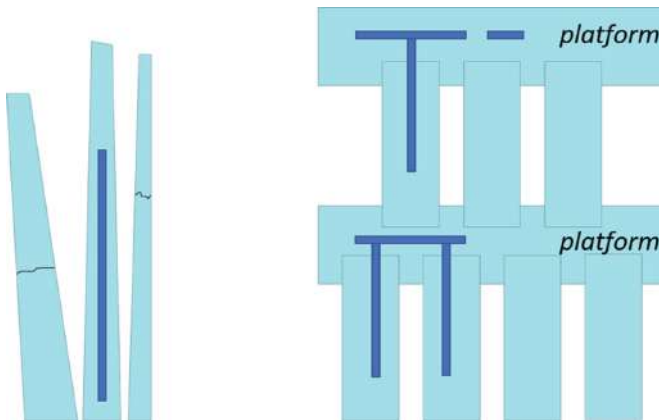


Fig. 6. Schematic representation of competency profiles, applicable to ICT professionals. I-profiles represent ‘do it yourself’ end-to-end professionals from the early days and the many specialist that today operate in a specific expertise field. Increasingly, platform integration profiles (T-shape) and cross-domain integration profiles (Π -shape) are required to build the IoT.

It requires the development and maintenance of technical and non-technical skills to be successful in:

- Solving societal challenges
- Addressing non-functional system aspects crucial for trust
 - Educating and involving society
 - Providing end users with options and personal data ownership
- Managing convergence on application level
 - Internet, IoT, mobile, IIoT, OT, data analytics, systems of systems, SDN
 - Cybersecurity
 - Cloud-, fog-, edge-computing and their distribution
 - Real-time, mission critical SW for specific application domains
 - VR, AR, robotics, digital twins, BIM...
- Interoperability and portability of functions and data
- Designing and deploying distributed Systems Architectures

But this also provides many development opportunities for IT professionals and an increase of jobs.

7 The Alliance for Internet of Things Innovation (AIOTI)

The Alliance for Internet of Things Innovation (www.aioti.eu) was kickstarted in 2015 and formally established in 2017 with the aim to address the cross functional and integration aspects of building a successful IoT. It is a member driven organisation with representatives from industry (large and small), academia and society that:

- Is at the forefront of IoT adoption, able to identify what is required to drive this adoption;
- Strives to break down silos so that the market for IoT can develop;
- Develops IoT ecosystem across vertical silos including start-ups and SMEs;
- Contributes to Large Scale Pilots to foster experimentation, replication and deployment;
- Supports convergence & interoperability of IoT standards;
- Gathers evidence on market obstacles for IoT deployment in a Digital Single Market.

by

- Promoting an integrative approach;
- Leveraging existing initiatives, be the missing link;
- Co-operating with other global regions while European values, including privacy and consumer protection, are maintained.

AIOTI embraces diversity, expressing the different views of interest group along the value chain. AIOTI is leveraging a structure of horizontal working groups, addressing common elements in technology research, ecosystems, standards and policies with an implementation driven approach in vertical working groups. This is depicted in Fig. 7.



Fig. 7. Working structure of the Alliance for Internet of Things Innovation (2018), combining horizontal and vertical working groups (WG)

8 Conclusions

We have taken an integral perspective on the development of IoT, beyond technology and clearly the Internet of Things holds many *promises*. But just as much as the promises, the development of IoT has many *challenges* requiring a new approach involving:

- Creating platforms by a strong collaborative approach beyond technology;
- Socio-economic aspects in a Human Centric IoT;
- The education and involvement of end-users;
- Privacy, security, resilience... and many more non-functional aspects;
- Critical architectural choices;
- Real scale experimentation.

Those elements and technical elements should be addressed in an integrated approach, on the one hand leveraging specialist companies and individuals, separating concerns, but linking them in an overall approach. The Alliance for Internet of Things Innovation promotes and drives such approach.

Interesting and responsible opportunities emerge for IT professionals, playing key roles in architecture and platform integration and in linking application domain specialists and end-users.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





IoT: Do We Have a Choice?

Draft IFIP Position Paper

Leon Strous^(✉) and IFIP Domain Committee on IoT

Amsterdam, The Netherlands
strous@iae.nl

Abstract. Many experts and organizations are addressing the Internet of Things (IoT) in policy statements, papers and conferences. There are many aspects to be considered when talking about IoT. The International Federation for Information Processing (IFIP) contributes to the discussion by investigating what choices can or must be made regarding these various aspects. And by addressing the question what choices various stakeholders should have. This draft version of the position paper was discussed at the IFIP IoT working conference on 18-09-2018 and in the IFIP General Assembly on 23-09-2018. The outcome of these discussions will be processed in the final version of the position paper.

Keywords: Position paper · Choices · IoT dimensions · IoT lifecycle

1 Introduction

IoT is hot. Many experts and organizations are addressing the topic in policy statements, papers and conferences. There are many aspects to be looked at when talking about IoT. IFIP wants to contribute to the discussion by investigating what choices can or must be made regarding these various aspects. And by addressing the question what choices various stakeholders should have. This paper briefly lists the aspects and dimensions of the IoT. Then IFIP's position on some major questions and choices is presented. It concludes with an overview of (possible) contributions that are already made or can be made by IFIP and its member societies and by ICT professionals in general to the open questions.

This draft version of the position paper was discussed at the IFIP IoT working conference on 18-09-2018 and in the IFIP General Assembly on 23-09-2018. The outcome of these discussions will be processed in the final version of the position paper.

2 Definition

More than one definition of the IoT exists. For the essence of this paper the definition is not the most important element. It was therefore decided to neither choose one from the list nor create one of our own to guide the discussion. A few examples of definitions are

added in an annex at the end of the paper. It is important to note however that while the IoT can be seen as a global infrastructure several vertical domains for distinct applications can be defined.

3 Aspects and Dimensions

3.1 Opportunities Versus Threats

Like every new technology also the Internet of Things offers opportunities for progress and application for beneficial purposes while at the same time it introduces or increases risks and threats. When addressing choices about various aspects, both sides will be considered.

3.2 Dimensions

As mentioned in the introduction, there are many aspects to be looked at when talking about IoT and discussing what choices can or must be made regarding these various aspects. In the current literature, many lists of aspects are a mixture of types of aspects. In an attempt to structure this, a three-dimensional model is proposed. The three dimensions would distinguish choices to be made:

- by whom (individuals and organizations)
- during which phase of the lifecycle of an IoT application
- about which issues.

By Whom

Choices are to be made by individuals and by organizations. An individual can be in the role of ICT professional developing IoT infrastructure or IoT applications or in the role of a user of IoT. Organizations can be in the role of user, of ICT industry developing IoT hardware and software or of authorities/regulators responsible for policies, standardization, legislation and other types of regulation.

Phases of an IoT lifecycle

Many lifecycle phases of products, systems and applications can be found in literature. Generally speaking the following phases can also be distinguished for an IoT application:

- Analysis/design;
- Development/production;
- Operation/maintenance;
- Disposition.

Issues

The broad spectrum of issues to be considered includes:

- Technical issues
- Legal issues (including liability)

- Ethical issues
- Education, training, awareness
- Usability and accessibility issues/freedom of choice and personalization issues
- Environmental issues
- Privacy issues
- Risks, Security, Resilience
- Impact on persons and society
- Professionalism/duty of care.

4 IFIP's Position on Major Questions and Choices

This position paper is not a series of positions on the technologies in the Internet of Things but it is a series of statements about choices that can be made and/or should be made and that should be enabled by technologies and/or policies. As a federation of societies of ICT professionals, for our positions we take the perspective of a human centred IoT: *“A human centred IoT would imply an environment where IoT will empower people and not transform them into hostages of technology”* [1].

The most elementary choice is the question “can I choose not to use an IoT?”. The answer to this question is not straightforward for all cases. There may be arguments e.g. for national security or environmental reasons to limit the choices. In the following paragraphs this and a number of other questions will be addressed. In each paragraph IFIP's position on a variety of aspects is presented and substantiated.

The paragraphs are following the dimension “By whom” (see Sect. 3.2).

4.1 ICT Professional

1. **IFIP's position is that an ICT professional should have sufficient professional and ethical competencies to make the right choices when designing, developing, implementing, operating or managing software/hardware as part of an Internet of Things that is able to offer choices to its' users.**

Having sufficient professional and ethical competencies is a general requirement for ICT professionals. However, in an IoT environment this is especially important because users may not be aware of the fact that choices are, could or should be possible. Users also may not be in a position to demand choices or to influence the usage of collected data. Therefore, the professionals should see to it that such choices are embedded and offered. The constraint of course is that also an ICT professional may not be in a position to decide upon the design etcetera. This means that a condition for making this work is to have professional and ethical competencies not only embedded in the codes of ethics of societies of professionals but also in companies' policies. And to have a work environment that is supportive of putting these policies into practice.

II. IFIP's position is that ICT professionals have a choice to educate/inform users on both the potential benefits and the risks of the Internet of Things the users are confronted with.

Users should be informed about the benefits and risks of Internet of Things applications they use. If the owner/developer of such applications does not (sufficiently) inform the users, ICT professionals have a choice, maybe even a duty, to do this, for instance via research papers and publications. In order to be able to do this, there should be no legal liability when publishing such results.

4.2 User

Both individuals and organizations can be in the role of user.

- III. IFIP's position is that users at least must have a choice to switch off the connection/not use the smart part of smart devices. In other words, users should have an opt in or opt out choice.**
- IV. IFIP's position is that it supports the possibility to empower users in such a way that they can control and personalize the behaviour of smart objects and associated applications through appropriate design tools even if they do not have programming knowledge.**

For example, if a smart meter gives the energy company full insight in the user's energy consumption, the user should have the option to not provide this information. This means that policies/regulations/legislations should allow for this and also the technology/devices should make this possible. Users should be aware of the consequences of both the opt in and opt out choice.

There may be applications or circumstances where it is not possible or desirable to give users an opt in or opt out choice, for instance in cases where national security is at stake. When this is the case, it should be clearly explained to users.

- V. IFIP's position is that users should inform themselves about the various aspects (benefits/risks) of the devices that are connected in the IoT they are using.**

While ICT professionals and ICT industry have a choice, or actually an obligation, to educate/inform users, these users have a choice, or also perhaps an obligation, to inform themselves. This can be by simply reading the information provided or asking for information if that is not provided. A condition to help users is the availability of "a set of the right questions".

- VI. IFIP's position is that involving users in the design/development of IoT (application) should be encouraged.**

Users are not only passive users but are also often people who possess knowledge and can contribute in the design/development of IoT. Having a say – if possible, in the design process – would be one way to make them more active.

4.3 ICT Industry

VII. IFIP's position is that the ICT industry providing IoT applications should inform users about the benefits and potential risks.

This should not be a choice but an obligation. It has to be clear for users for which purposes data are collected. A mechanism needs to be in place to assure the security and protection of such collected data and providers should inform users about these mechanisms. It should also be made clear what the consequences of either choice (opt in or opt out) are.

VIII. IFIP's position is that the ICT industry should not develop IoT applications that provide data that can be used without the owners of the data knowing about the use or consenting to it. The ICT industry has a choice not to do this.

Owners of data, both personal data or data that can be linked to persons in an indirect way, should know who is doing what with their data and they should have the right to give consent for such usage. This may not be possible in all cases but that should then also be clear.

4.4 Authority/Regulator

IX. IFIP's position is that policymakers/regulators should take into account the interests of users when regulating the use of (personal) data (including data that can be linked to a person in an indirect way e.g. via home, car, etc.).

Policymakers/regulators have a choice to balance the interests of various stakeholders in the applications and their data. It is important that policies and regulations provide the conditions for the choices that users and providers can or should be able to make.

5 Possible Actions

IFIP, its member societies and their members can contribute to solve the "choice problems" addressed in the previous chapter. What can be done:

- Check/promote the presence of the professional and ethical competencies issue, for example in codes of ethics of professional societies and in companies' HR policies.
- Provide a "set of the right questions".
- Promote the position statements to the professionals, users, industry and authorities.
- Research the benefits and risks of the various Internet of Things applications.
- Increase research of those aspects that are insufficiently addressed and/or that are gaining more and more importance. Examples could be: (1) With the increasing number of IoT devices will there be energy to run all of them? IoT is requesting the production of low power devices, that means the use of optimisation techniques, and the direction is to have dedicated devices to each need or function. (2) With the increasing number of IoT applications, ethical (privacy, surveillance etc.) and

security issues are becoming more and more important due to the use, design and implementation of such applications.

6 Annex. Sample Definitions

ITU [2]

“Internet of Things (IoT): A global infrastructure for the information society, enabling advanced services by interconnecting physical and virtual things based on existing and evolving interoperable information and communication technologies.

Thing: an object of the physical world or the information world, which is capable of being identified and integrated into communication networks.

Device: a piece of equipment with the mandatory capability of communication and the optional capabilities of sensing, actuation, data capture, data storage, and data processing.”

Wikipedia [3]

“The **Internet of Things (IoT)** is the inter-networking of physical devices (also referred to as “thing”, “object”, “connected devices” or “smart devices”) such as vehicles, buildings, and other items embedded with electronics, software, sensors, actuators, and network connectivity which enable these objects to collect and exchange data. The IoT allows objects and their environments to be sensed or controlled remotely across existing network infrastructure, creating opportunities for more direct integration of the physical world into computer-based systems, and resulting in improved efficiency, accuracy and economic benefit in addition to reduced human intervention. When IoT is augmented with sensors and actuators, the technology becomes an instance of the more general class of cyber-physical systems, which also encompasses technologies such as smart grids, virtual power plants, smart homes, intelligent transportation and smart cities. Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing Internet infrastructure.”

Gubbi et al. [4]

“The worldwide network of interconnected objects uniquely addressable based on standard communication protocols.” and “Our definition of the Internet of Things for smart environments is: Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless ubiquitous sensing, data analytics and information representation with Cloud computing as the unifying framework.”

References

1. EC Staff working document “Advancing the Internet of Things in Europe”, April 2016. <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-advancing-internet-things-europe>
2. ITU Recommendation Y.4000/Y.2060 - Overview of the Internet of things. <https://www.itu.int/rec/T-REC-Y.2060-201206-I>
3. https://en.wikipedia.org/wiki/Internet_of_things. Accessed 14 Aug 2017, 21:22 Amsterdam, (Slightly Adjusted)
4. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **29**(7), 1645–1660 (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



IFIPIoT 2018 Refereed Papers



The Outcomes of the Implementation of Internet of Things: A Public Value Perspective

Ott Velsberg^(✉)

Department of Informatics, Umeå University, Umeå, Sweden
ott.velsberg@umu.se

Abstract. In recent years, Internet of Things (IoT) has gained presence in all areas of life. Whilst private sector is the forerunner in the adoption of these devices, public sector usage has been lagging. With the rise of smart initiatives, public organizations are increasingly implementing IoT. The necessity to know in which areas of public sector IoT has been implemented and what public value has been derived, has gained importance as reporting of the cost efficiency and benefits of these initiatives has gained attention. This paper seeks to determine the importance of IoT in the public sector from the perspective of public value. IoT implementation in Estonian municipalities was studied to provide concrete data on the use of IoT. Next to efficiency, which is a known goal of IT implementation, the research findings suggest that while IoT has generated public value, there has been a shift in value creation with other outcomes such as effectiveness, transparency and collaboration gaining increased presence. While IoT shows great promise for public value creation, more research is needed to study how public sector can leverage these devices to harvest more benefits than the simple automatization of work processes.

Keywords: Public values · Public sector · Value creation · Internet of Things

1 Introduction

The term Internet of Things (IoT) has attracted considerable attention in academics, industry and public sector, and is envisioned as a global network of machines and devices that can interact with each other [23]. The IoT has been regarded as one of the most important areas of future technology [23] and to be at the core of the fourth industrial revolution [19]. The European Commission envisions IoT as an indispensable component towards the digitisation of our society and economy [8]. It is forecasted that the market value of IoT in the EU will exceed one trillion euros by 2020 [7], and by 2025 the yearly global economic impact of IoT is estimated to be between 3 trillion to 9 trillion euros, accounting close to 11% of the world economy [25].

Research indicates that by 2020, close to 26% of all IoT solutions will be consumed by public sector [12]. A high rate of expenditure can be predicted, as national and local governments have been voracious consumers of information technologies (IT) in the past [29]. However, recent studies indicate that close to 60% of IoT initiatives stall at the Proof of Concept stage and 75% of completed IoT projects fail to produce value

[16]. With regards to this, it becomes questionable whether the value created from IoT overweighs the financial risks taken by the public sector.

Creating public value has been a key focus in digitizing public sector organizations. If the public sector continues to invest in IoT, our innate logic tells us that it must create public value. Agarwal and Lucas [1] argued that to contribute to the IS field it is fundamental to demonstrate the value of IT. Following the same argumentation, it is crucial to demonstrate the public value of IoT. However, to date, the research on IoT and public value creation has been scarce.

The aim of this research is to offer a contribution regarding the public value created by IoT, asking: What public value does IoT create in the public sector? The study offers a glimpse on the main values created by IoT, based on an extensive qualitative study on 46 municipalities in Estonia. The paper briefly introduces the concept of public value and IoT, including an overview of IoT implementation in the public sector.

The paper is structured as follows. First, the theoretical basis is set by introducing the concept of public value. Next, the concept of IoT and its role in public sector will be introduced. In the following parts, the research design and findings are presented. This will be followed by a discussion of findings and a conclusion.

2 Theoretical Background

This section introduces the concept of public value, which is regarded as one of the main reasons for public sector innovation. The following section introduces the concept of IoT and discuss its importance in generating public value.

2.1 Public Value Perspective

Public entities such as governments, municipalities and county councils are constantly seeking to address contemporary challenges and opportunities through utilizing emerging technologies [29]. While there are many motivators for adopting IT solutions, a primary goal of IT initiatives is to create and deliver public value [18]. Public value focuses on governmentally produced benefits that serve the interests of stakeholders both inside and outside the public organization [14]. The creation of public value entails balancing competing public interests through emphasizing collective preferences and expectations of government, citizens and other stakeholders who consume the services [5, 27]. Consequently, public sector bases its decisions to implement IT on the expected benefits and conflicting demands of various stakeholders [30].

It is widely acknowledged that IT creates public value. A successful IT implementation is believed to drastically change how governmental organizations operate, bringing forth substantial organizational, technical, business and societal benefits [10, 20]. IT-based value manifests itself both in terms of economic values e.g. operational efficiency, and non-economic values e.g. trust in government and sustainability [29]. For the associated stakeholders, e.g. citizens, users, public administrators and politicians, the public value can manifest differently, e.g. administrative personnel and managers from public sector might place high value on accountability, c.f. [10], while

citizens might value accessibility to governmental information [38]. Harrison et al. [14] divided public value into seven categories, namely:

1. Economic – impacts on current or future income, asset values, liabilities, entitlements, or other aspects of wealth or risks to any of the above.
2. Political – impacts on a person’s or group’s influence on government actions or policy, or their role in political affairs, influence in political parties or prospects for public office.
3. Social – impacts on family or community relationships, social mobility, status, and identity.
4. Strategic – impacts on person’s or group’s economic or political advantage or opportunities, goals, and resources for innovation or planning.
5. Quality of life – impacts on individual and household health, security, satisfaction, and general well-being.
6. Ideological – impacts on beliefs, moral or ethical commitments, alignment of government actions/policies or social outcomes with beliefs, or moral or ethical positions.
7. Stewardship – impacts on the public’s view of government officials as faithful steward.

The incorporation of public value in the discussion of public sector investment in IT is especially relevant as it allows to monitor the outcome of government investment and understand the relationship between IT and public value delivery [36]. Understanding the relationship between IT and public value creation can assist public organizations in the pursuit of suitable use of technology to benefit society and secure the cost efficiency of public value creation [3].

2.2 Role of the Internet of Things

The IoT has rapidly gained presence and is sometimes regarded as the most disruptive phase of the Internet revolution [2, 34]. IoT is ubiquitous by nature and is present in different almost identical concepts, such as “Internet of Everything”, “ubiquitous computing”, “pervasive computing” and “ambient intelligence”, whereas the differences between the terms are of academic nature [9]. IoT is a general term used for objects interconnected through networks, that encompass processing and sensor capabilities, allowing the devices to transmit recorded information from the outside environment. IoT allows information, resources, “things”, e.g. sensors, beacons, actuators, mobile phones etc., to interact with each other and cooperate with their smart components to reach common goals [15]. Realizing the value of IoT requires integration of IT infrastructures and information services - such as RFID tags, wireless broadband and geographic information systems [33]. Sensory data must be gathered from distributed smart objects and be transmitted using a communication infrastructure, which can encompass both wired and wireless communication technologies [32]. IoT allows to link real-life objects with the virtual world - providing anytime, anyplace connectivity for anything [34].

These capabilities can be applied to everyday objects, thus affecting all areas of life [9]. IoT will change the way we collect, analyse, and respond to data, creating opportunities for

individuals, governments, and businesses to develop new business models and forms of interaction that take advantage of ubiquitous computing power embedded in objects [13]. While the majority of IoT initiatives have been implemented in the private sector, studies have indicated that the importance of IoT in public sector is increasing [4, 12]. This is driven by the rapid development of IT, and expanding efforts by national and local governments to change how they operate. Local governments are increasingly turning towards new information systems, often utilizing IoT, to develop livable, economically sustainable and efficient living environments. IoT is implemented in various settings from cities to rural areas. Utilizing IoT devices allows local governments to monitor and take immediate action on almost every aspect of urban and rural space, and provide citizens with relevant information and services [33, 35]. Among others, IoT can be used to create new and enhance existing services, improve efficiency and effectiveness of internal management and service delivery, and foster collaboration with different stakeholders [9, 22]. Due to the dynamic nature of public sector, the requirements for IoT devices can vary significantly. The IoT devices that can swiftly and accurately transmit information can be crucial for winter road maintenance, while speed would not be a requirement for water consumption monitoring.

In the context of public sector, IoT has the potential to improve several areas, including: healthcare, education, utilities, infrastructure, buildings, environment and culture. With regards to this, IoT becomes pivotal in streamlining governmental processes and engaging citizens in all areas of local governance. However, to date, little is known on how IoT generates public value. Consequently, there also lacks knowledge on how local governments could benefit from implementing IoT.

3 Public Value Framework

For the study in hand, it was considered suitable to use the six public value generating mechanisms proposed by Harrison et al. [14] to understand what public value IoT creates in public sector (See Table 1). Harrison et al. [14] created the set of value generators to specify how public value is created, i.e. which government action lead to the creation of public value. For analysis, the six public value mechanisms proposed by Harrison et al. [14] were used for guidance. To investigate the relationship between IoT and public value generating mechanisms, a public value framework was created to assess the IoT initiatives and to distinguish between the outcomes (See Table 1). Based on Nam and Pardo [28] the evaluation dimensions were presented from the external view (interactions with citizens, private companies and other relevant non-governmental actors) and from the internal view (interactions within the local government). Identifying and measuring the public value of IoT is complicated as different stakeholders may have different attitudes on what is regarded a successful outcome of an IT initiative [37]. As a result, public value creation was studied from the perspective of public organizations. Similarly to this study, most research on IT value studies the outcome of past IT investments through a post hoc analysis [21]. As the study is concerned with how IoT solutions add value to local governments and public, broader economic factors are not considered unless they directly relate to governmental or public impact.

Table 1. Public value framework.

Dimensions	Management (internal view)	Service delivery (external view)
Efficiency	Concentrates on the internal managerial efficiency in terms of obtaining increased outputs, workloads, activities, processes and goal attainment	Concentrates on the efficiency of producing and delivering services
Effectiveness	Concentrates on the quality of internal management	Concentrates on the quality of services delivered
Intrinsic enhancements	Concentrates on the changing environment or circumstances for governmental stakeholders	Concentrates on the changing environment or circumstances for non-governmental parties
Transparency	Concentrates on the access of information or processes inside the local government	Concentrates on the external access to information and processes regarding service provision and delivery
Participation	Concentrates on the frequency and intensity of direct involvement of internal stakeholders in decision making or operation of government	Concentrates on the frequency and intensity of direct involvement of external stakeholders in decision making or operation of government
Collaboration	Concentrates on the improvement of collaboration inside the local government in terms of sharing responsibility or authority for governmental processes and actions	Concentrates on the collaboration between governmental and non-governmental parties in terms of sharing responsibility or authority for governmental processes and actions

4 Research Design

To understand how IoT creates public value, a qualitative study was carried out on the use of IoT in Estonian municipalities. The following section provides information on the research domain, data collection and analysis.

4.1 Data Collection

To paint a thorough picture of the use of IoT in public sector, the research concentrated on the use of IoT in Estonian municipalities. The data was collected between March and December 2017 through semi-structured interviews and supporting documentation reviews.

The research participants were found by first contacting all 202 municipalities in Estonia via e-mail regarding the use of IoT in public sector, the municipalities that had implemented IoT solutions were included to the study and the municipalities who were in the planning phase of IoT implementation or who had not implemented IoT solutions were excluded. Thus, ex ante descriptions of the predicted outcome of IoT were not included in the study. In total 81 municipalities replied, of whom 46 were included in

the study. From those 46 municipalities, 67 participants were contacted and interviewed regarding the implemented IoT solutions. In some municipalities, more than one public official was interviewed due to a lack of involvement of the interviewee with other IoT solutions.

The interviews were conducted in person where possible, and over the phone in other circumstances. An interview guide was used to determine the course of the interviews, which included open-ended questions that allowed to include topics that were not predetermined by the interview guide. The type of questions in the interview guide included the type of solutions implemented, the outcomes of the solutions, collaboration with other stakeholders, the role of IoT, measurement techniques for the outcomes etc. In Table 2 an overview of the interview questions is presented. The interview guides were iteratively changed, according to the research conducted prior to the interviews - publicly available information on each municipality was gathered to determine the type of solutions introduced.

Table 2. Overview of interview questions.

Interview question examples:
What IoT solutions have been implemented?
What has been the guiding principle behind the deployment of IoT?
What benefits has the IoT brought?
How has the IoT affected municipal processes?
How do you verify the integrity of the collected data?
Why did the municipality decide to use the devices?
Would it be financially beneficial for other municipalities to use these devices?

The interviews lasted between 20 and 120 min, with an average of 57 min. All interviews were audio recorded and later transcribed. The participants had been using IoT solutions from five months to five years - providing enough time to realize the value of IoT initiatives [24]. A summary of all interviews will be presented in Table 3.

Table 3. Summary of interviews.

Public sector representatives	Total number of respondents
Chief information officer	7
Environmental and municipal advisor	3
Head of economics department	9
Head of public administrative unit	4
Head of road maintenance	1
Head of utilities	7
Municipal mayor	14
Municipal vice-mayor	22

Supporting documentation such as procurement documents, public documents, and technical files were collected from governmental databases and online sources, e.g. municipal websites and from technology providers. Technical documentation was attained on both the system and the device where possible. Procurement documents which allowed to better understand the reasons and granularities of the implementation of any device were studied where possible. Public documents used included governmental reports, public statements etc. Secondary data sources were used to prepare for interviews, map government priorities etc. A summary of all additional procedures for data collection are presented in Table 4.

Table 4. Summary of data collection.

Supporting documentation	
Procurement documents	118
Public documents	173
Technical documents	131

4.2 Data Analysis

To analyse the data, the recommended procedures for qualitative research were followed to guide through the three steps of coding [6, 26]. During the first round of coding an initial coding scheme was developed. As the research was concerned with identifying what public value IoT generates, the coding scheme was based on public value framework (See Table 1). During the second round of coding, examples of IoT generated public value were identified. As the study focused on how IoT has generated public value for the municipalities, the perspective of citizens and other stakeholders were considered through municipal perspective. The different dimensions were recorded altogether 2644 times. From the mentioned segments, 1481 concentrated on the external view. The remaining segments were divided between the internal dimensions. A summary of coded segments can be seen in Table 5.

Table 5. Summary of coded segments.

Dimensions	Management (internal view)	Service delivery (external view)
Efficiency	431	339
Effectiveness	298	405
Intrinsic enhancements	67	74
Transparency	113	301
Participation	143	185
Collaboration	111	177
Total	1163	1481

The third round of coding compared previously coded segments to summarize shared features and variance within and across research sites. Multiple data sources allowed to compare, contrast and triangulate data [26]. ATLAS.ti was used throughout the study for data analysis. Findings are presented in the following section.

5 Empirical Analysis

In recent years, IoT has gained traction in public discourse as IoT solutions are increasingly implemented. This research concentrates on the use of IoT in Estonian municipalities - studying the public value created by these solutions.

5.1 Mapping the Estonian IoT Solutions

IoT solutions can be implemented for virtually any purpose, however the study identified some core areas in which the implementation of IoT was more common, namely buildings and transportation. In total, 158 IoT solutions were implemented, from those, the study identified 30 different IoT solution types. Table 6 will present an overview of the implemented solutions and the number of municipalities that implemented those.

Table 6. Summary of IoT solutions.

Area	Number of solutions (percentage of all solutions)	Overview of the solutions	Examples (number of solutions)
Buildings	29 (18.3%)	Systems that include simple motion sensors to regulate lightning to more complex systems that regulate window canopies depending on lightning and temperature	Heating system (9); Smart lightning (10); Ventilation system (8); Window canopies (2)
Infrastructure	21 (13.3%)	Systems that include simple motion sensors to turn on and regulate lightning, to more complex systems that regulate lightning depending on vehicle speed and depending on public transportation movement. Pothole identification system allows to identify and map the condition of roads	Pothole identification system (1); Intelligent lightning (20)
Healthcare	2 (1.3%)	Sensors collecting physical vitals and transmitting them to healthcare professionals and public officials. Allows to notify of emergencies	SOS-bracelets (2)
Security	19 (12%)	Vehicle surveillance systems collect, analyze and transmit information on	Smoke detectors (3);

(continued)

Table 6. (continued)

Area	Number of solutions (percentage of all solutions)	Overview of the solutions	Examples (number of solutions)
		vehicles. Smoke detectors and surveillance systems allow remote access and improved detection	Surveillance system (7); Vehicle surveillance system (9)
Transportation	49 (31%)	Connected buses and fleet telematics allow remote monitoring of location, performance and behavior of vehicles Remote passenger validation system allows to validate passengers and track their movements Traffic and pedestrian/cyclist counter allows to count the number of individuals and various other elements, e.g. their speed and whether they wear bicycle helmets Smart parking allows to track and visualize the available parking spaces	Connected buses (5); Fleet telematics (34); Pedestrian/cyclist counter (2); Remote passenger validation system (1); Self-driving buses (1); Smart bicycle parking (3); Traffic counter (3)
Utilities	23 (14.6%)	Array of sensors that allow to improve the processes and automate the monitoring and maintenance of utility systems, e.g. geothermal systems, garbage transfer stations and wastewater treatment plants. Smart meters allow to collect and analyze water consumption and automate various processes, e.g. billing and detection of leakage	Biomass heating systems (5); Garbage transfer station (1); Garbage bins (2); Geothermal systems (6); Water meters (5); Water plant (2); Wastewater treatment plant (2)
Weather and environmental monitoring	15 (9.5%)	Array of sensors that collect and analyze environmental information, such as water and air quality, temperature and constituents. The system alerts when these values exceed a set threshold	Air quality sensors (3); Noise sensors (1); Pavement sensors (1); Snow monitoring (3); Water monitoring (3); Weather Stations (4)

Table 7. Number of solutions with identified public value.

Dimensions	Number of solutions		Total number of solutions
	Management (internal view)	Service delivery (external view)	
Efficiency	68	36	75
Effectiveness	64	53	94
Intrinsic enhancements	16	7	19
Transparency	28	64	67
Participation	24	17	32
Collaboration	38	42	45

Table 7 presents a summary of solutions where public value was identified. A solution could simultaneously provide public value in many dimensions, for example there were 75 solutions which generated efficiency, from those solutions, 68 produced public value for management and 36 solutions produced public value in service delivery.

The following part presents how IoT impacted the six public value mechanisms, drawing on specific IoT implementation examples for illustration, see Table 6.

5.2 Efficiency

The study identified that 64.5% of solutions (102 solutions) had efficiency as their primary desired outcome. From all solutions, 47.5% of solutions (75 solutions) identified efficiency as an achieved outcome. While reducing costs was the main consideration during the procurement, it was difficult for municipalities to isolate the effects and measure the outcome. Consequently, improvements in efficiency were visible in terms of productivity, most commonly through reduced time and improved communication. The devices allowed municipal employees to work everywhere at any time without physically being present. This allowed to get a better and timelier overview of work processes, coordinate and reduce the workload. The devices allowed to verify whether service contractors and systems worked as needed. Through automatic analysis it was possible to identify deviations that would have otherwise remained unnoticed. Even though IoT in some instances increased the hourly rate of work, the number of hours spent decreased. For instance, the fleet telematics enabled maintenance managers to re-route optimal vehicles, better maintain roads, and reduce the time vehicles were standing idle. Commonly implemented solutions included fleet telematics which helped to resolve disputes among stakeholders and reduced strain on municipalities to provide timely information to citizens. At the forefront of cost cutting were solutions related to energy consumption reduction, however the use of motion and CO₂ sensors were difficult to utilize to their fullest potential due to the human element which interfered with the automatization, i.e. people opened windows instead of allowing the ventilation system to operate independently. To counter the human element, various approaches were taken, from employee training to changes in the

physical environment, i.e. making it impossible to open windows. Still, cost efficiency remained difficult to achieve and often even harder to measure due to the changing environmental elements, e.g. changing weather conditions made it hard to compare changes on yearly basis.

5.3 Effectiveness

Improved effectiveness was the most commonly achieved outcome from the use of IoT, identified as an achieved outcome in 59.5% of solutions (94 solutions). Although the implementation of IoT systems commonly had a goal to gain efficiency, especially in financial terms, most highly valued outcomes identified by the municipalities were derived from improved effectiveness. The devices allowed public officials to reduce unnecessary work (avoid personal interaction, provide easier access to information, allow for increased control, and offer convenience), and improve decision making and work outcomes (services could be personalized and provided in timely manner, investments were more targeted). Data provided by the sensors allowed to evaluate the work of devices, e.g. heating systems and private contractors, e.g. garbage disposal providers. In case of deviations, municipalities were notified of the changes. This allowed to improve the quality of service delivery and through that reduce costs as it took less time to achieve the intended outcome. Controversially, the devices often proved to be more efficient in identifying the shortcomings in effectiveness rather than aiding in improving them. For example, the pothole identifying system allowed to map the problematic roads, however the information did not necessarily lead to improvements as the processes surrounding the road management were not improved to reflect the capabilities of the data generated by IoT. As an example, a municipality used the system to identify potholes, but made no changes to organizational operation and road management strategy.

5.4 Intrinsic Enhancements

Intrinsic enhancements were identified as the least important outcome from the use of IoT, only identified in 12% of solutions (19 solutions) as an achieved outcome. In majority of cases intrinsic enhancements were not considered during the implementation, but rather emerged as unexpected outcomes during the use of the devices. Intrinsic enhancements considered here, i.e. less redundant work practices regarding communication, were by nature silent outcomes, as they were experienced by all, but the highest improvements would be expected at service level and for consumers, but not necessarily by service providers - in this case municipalities. This in turn contributed to the low priority of intrinsic enhancements.

5.5 Transparency

Improved transparency was identified as an achieved goal of IoT in 42.4% of solutions (67 solutions). IoT allowed municipalities and their employees to defend their decision making, overcome false accusations (reduce the impact of political loyalty), and helped to avoid unethical and unlawful actions, i.e. provide equal treatment for all stakeholders

and protect individual rights. Through making information available for the public, municipalities could reduce their workload, make more calculated decisions, foster trust and accountability, and receive valuable feedback on their services. When citizens demanded information, the public official could rely on IoT generated data to provide up-to-date information. Municipalities could further control whether service contractors worked in an agreed upon amount and according to contracts - helping to contribute to reliability and stability of service provision. This allowed to avoid conflicts between the municipality and the service contractors, paved the way towards a trusting relationship and enhanced cooperation. However, transparency remained the most controversial outcome of IoT implementation, as improvements in transparency generated countless unforeseen difficulties that resulted in reversing the improved work practices to the previous mode of operation. To illustrate, citizens started to ambush winter road maintenance vehicles when the information of the vehicle movement was made available online. This resulted in an overall stigma for improvements in transparency, as failures in municipal management could have given a political disadvantage in upcoming elections, or provided a competitive advantage for neighboring municipalities.

5.6 Participation

Improved participation was identified as an outcome of IoT implementation in 20.3% of solutions (32 solutions). Improved participation was never the primary intent of IoT implementation, however municipalities experienced improved participation both from local government and from external stakeholders. From the improvements in the dimension of participation, 24 solutions identified improvements on an internal level, improvements on an external level were visible in 17 solutions. However, silos existed whereby internal actors from different departments were not invested in the improvement of government operation even if capabilities to support it were created. Circumstances where internal stakeholders were involved consisted of instances where the system directly affected them in terms of their private or working life. While external stakeholders were more highly interested in decision making and in the operation of government, it was generally not welcomed by the government officials. Instead the generated data was used to support the sole decisions of managers involved with the IoT implementation, leaving out the external actors - mistrust and uncooperative behaviors remained wide-spread.

5.7 Collaboration

Improved collaboration was identified as an achieved outcome in 28.5% of solutions (45 solutions). Collaboration was wide-spread both internally and externally. However, in majority of cases, collaboration was required for IoT implementation, but not necessarily fostered it. For instance, vehicle surveillance systems required police involvement due to the data protection acts. Collaboration was often present before the implementation of IoT, however IoT greatly enhanced the level of involvement and supported influx of more relevant information. For instance, when municipalities noticed that the quality of work was consistently worse for some drivers, e.g. garbage

removal, they notified the private contractors, which allowed to improve the service delivery. Due to the availability of information, collaboration emerged between public officials and private contractors. For instance, when a technician previously had to be present to fix a biomass heating system, then following IoT implementation public officials took increasingly care of the maintenance. Similarly, when there were problems with private contractors' work, public officials directly contacted the people responsible. IoT provided capabilities that made municipalities adopt a more active role in service delivery. Indeed, municipalities had started to consider providing services without public-private collaboration. For instance, in winter road maintenance IoT created a situation where the public official became more actively invested in the maintenance than the private maintenance manager. Citizens on the other hand were involved in government operation through actively evaluating the available information, notifying of shortcomings. When services did not meet citizen expectations or citizens needed a custom service, they contacted for improvements. While citizens were involved, their participation in the governance was not wide-spread. With regards to this, citizens had limited impact, i.e. did not contribute, on the functioning of government.

6 Discussion

Previous research indicates that efficiency is typically the primary goal of IT initiatives [11], which also holds proof here. However, the findings suggest that public organizations are increasingly focusing their attention on other dimensions besides efficiency, most commonly on effectiveness and transparency. This goes against the typical notion of IoT which is mainly articulated in terms of economic value [31]. Municipalities included in the study often considered other dimensions as the most sought after, neglecting efficiency as an important outcome even if an element of efficiency, e.g. time or cost, was stated as the primary implementation goal. By contrast other elements, most notably effectiveness and transparency, were increasingly derived as outcomes of these government initiatives. With regards to this, IoT allows to create public value beyond efficiency.

The findings suggest that while collaboration, transparency and participation generated public value, they should not be viewed as an administrative goal [14]. The opposing needs of stakeholders could create disturbances that could negatively affect the effectiveness and efficiency of public organizations. Harrison et al. [14] argued that transparency, participation and collaboration should be viewed as means toward desirable ends. The findings of this study suggest that without clear goal, public organizations should refrain from making information publically available. Controlled secrecy in some instances could prove more beneficial for deriving public value. For instance, available information was used to interfere with service provision and how the system operated, e.g. stopping winter road maintenance vehicles or opening windows to regulate temperature. Hence, each case should be closely evaluated to determine whether external and/or internal actors should be engaged in the operation of government. While it is easier to continue in secrecy, transparency and collaboration had a strong effect on achieving public values from other dimensions, e.g. effectiveness and

efficiency, which otherwise were not achieved. Thus, government initiatives should be directed to stakeholders that allow to achieve the optimum public value.

The study illustrates that while some solutions are more beneficial than others, deriving public value from IoT initiatives requires public organizations to use the generated data to change how they operate. While the most substantial changes require IoT technology, management and policy changes make the biggest difference in the derived outcome of IoT. Using smart garbage bins or fleet telematics without changing how public organizations operate brought forward positive changes, but substantial benefits, required a change in management and policy. Public organizations had to change how they operate through using the devices and generated data in a unique way. For instance, they could continue to use fleet telematics for surveillance of vehicles, however when incorporating the generated data in investments, decision making or fostering closer collaboration with private organizations, the most substantial benefits occurred. As a result, the findings suggest that most public organizations are handling post-implementation phase of IoT inadequately, most notably regarding what to do with the produced data.

The research indicates that public organizations are not evaluating the outcome of IoT implementation. Throughout the study, rather few organizations had performed a formal evaluation of an IoT initiative. Similarly, the study by Jones and Hughes [17] identified that public organizations tended to rely on an assessment whether the technology works, rather than considering the social impact or value of the technology. If the outcome was evaluated, it was done through the prism of efficiency in an abstract manner. Without the formal assessment, it becomes difficult for public organizations to demonstrate the public value of IoT. Relying on abstract measurement systems, such as the number of complaints by citizens, could be a viable predictor of the success of an initiative, but it neglects the more substantial outcomes the technology could produce. There thus remains a need to explore how IoT created public value can be measured and is measured by public organizations. As the previous findings suggest, other aspects besides efficiency are gaining traction in IoT generated public value, the study concludes that the analysis of IoT should also consider other value generators besides efficiency when evaluating the solution. Evaluation and presentation of the different forms of public value generated by IoT could allow public organizations to avoid public scrutiny and improve the utilization of IoT. Especially, as many IoT projects were sealed off from external stakeholders due to the fear of criticism and political harm.

7 Conclusion

To conclude, IoT has the capability to create public value and the intended public value dimensions have widened from the goal of improved efficiency. There has been a shift in value creation with other goals and outcomes such as effectiveness, transparency and collaboration gaining increased presence. Furthermore, the data suggest that the evaluation of IoT remains largely insufficient, and is mostly done abstractly or through financial metrics alone, which inhibits capturing the full potential of IoT. Public value derivation cannot rest on the implementation of IoT solutions alone, but must include

improvements in management and policy according to the ample data generated by the devices.

Finally, there are limitations to be acknowledged. First, the research studied public value through the perspective of public organizations, however public values are rarely identical for stakeholders. Thus, additional research is required to study public value from the perspective of stakeholders like citizens and private organizations. Secondly, evaluating public value is never an easy task and more longitudinal studies could offer further in-depth understanding of public value created by IoT.

References

1. Agarwal, R., Lucas, H.C.: The information systems identity crisis: focusing on high-visibility and high-impact research. *MIS Q.* **29**(3), 381–398 (2005)
2. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. *Comput. Netw.* **54**(15), 2787–2805 (2010)
3. Bannister, F., Connolly, R.: The trouble with transparency: a critical review of openness in e-government. *Policy Internet* **3**(1), 1–30 (2011)
4. Bradley, J., Reberger, C., Dixit, A., Gupta, V.: Internet of everything: a \$4.6 trillion public-sector opportunity. Cisco (2013)
5. Cordella, A., Bonina, C.M.: A public value perspective for ICT enabled public sector reforms: a theoretical reflection. *Gov. Inf. Q.* **29**(4), 512–520 (2012)
6. Eisenhardt, K.M.: Building theories from case study research. *Acad. Manag. Rev.* **14**(4), 532–550 (1989)
7. European Commission: Definition of a research and innovation policy leveraging cloud computing and IoT combination. Digital Single Market (2015)
8. European Commission: The Internet of Things. <https://ec.europa.eu/digital-single-market/en/policies/internet-things>. Accessed 01 Feb 2018
9. Friedewald, M., Raabe, O.: Ubiquitous computing: an overview of technology impacts. *Telemat. Inform.* **28**(2), 55–65 (2011)
10. Gil-Garcia, J.R., Chengalur-Smith, I., Duchessi, P.: Collaborative e-Government: impediments and benefits of information-sharing projects in the public sector. *Eur. J. Inf. Syst.* **16**(2), 121–133 (2007)
11. Gil-Garcia, J.R., Zhang, J., Puron-Cid, G.: Conceptualizing smartness in government: an integrative and multi-dimensional view. *Gov. Inf. Q.* **33**(3), 524–534 (2016)
12. GrowthEnabler: Market Pulse Report, Internet of Things (IoT) (2017)
13. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of things (IoT): a vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **29**(7), 1645–1660 (2013)
14. Harrison, T.M., et al.: Open government and e-government: democratic challenges from a public value perspective. *Inf. Polity* **17**(2), 83–97 (2012)
15. Hermann, M., Pentek, T., Otto, B.: Design principles for Industrie 4.0 scenarios. In: The Proceedings of the 49th Hawaii International Conference on System Sciences, Koloa, HI, USA (2016)
16. Johansen, C., Culp, B., Mora, M.: Cisco survey reveals close to three-fourths of IoT projects are failing, Cisco (2017)
17. Jones, S., Hughes, J.: Understanding IS evaluation as a complex social process: a case study of UK local authority. *Eur. J. Inf. Syst.* **10**(4), 189–203 (2001)

18. Jørgensen, T.B., Bozeman, B.: Public values: an inventory. *Adm. Soc.* **39**(3), 354–381 (2007)
19. Kagermann, H., Wahlster, W., Helbig, J.: Recommendations for implementing the strategic initiative Industrie 4.0. (2013)
20. Krishnamurthy, R., Desouza, K.C.: Tony Parham: fostering innovation DNA in the commonwealth of Massachusetts (2014)
21. Kohli, R., Grover, V.: Business value of IT: an essay on expanding research directions to keep up with the times. *J. Assoc. Inf. Syst.* **9**(2), 23–39 (2008)
22. Lee, J., Lee, H.: Developing and validating a citizen-centric typology for smart city services. *Gov. Inf. Q.* **31**(1), 93–105 (2014)
23. Lee, J., Lee, H.: The Internet of Things (IoT): applications, investments, and challenges for enterprises. *Bus. Horiz.* **58**(4), 431–440 (2015)
24. Marchand, D.A., Kettinger, W.J., Rollins, J.D.: Information orientation: people, technology and the bottom line. *Sloan Manag. Rev.* **41**(4), 69–80 (2000)
25. McKinsey: The Internet of Things: Mapping the Value Beyond the Hype. McKinsey Global Institute (2015)
26. Miles, M.B., Huberman, A.M., Saldana, J.: *Qualitative Data Analysis: A Methods Sourcebook*, 3rd edn. SAGE Publications, Thousand Oaks (2014)
27. Moore, M.H.: *Creating Public Value: Strategic Management in Government*. Harvard University Press, Cambridge (1995)
28. Nam, T., Pardo, T.A.: The changing face of a city government: a case study of Philly311. *Gov. Inf. Q.* **31**(1), 1–9 (2014)
29. Pang, M.S., Lee, G., DeLone, W.H.: IT resources, organizational capabilities, and value creation in public-sector organizations: a public-value management perspective. *J. Inf. Technol.* **29**(3), 187–205 (2014)
30. Pereira, G.V., Macadar, M.A., Luciano, E.M., Testa, M.G.: Delivering public value through open government data initiatives in a Smart City context. *Inf. Syst. Front.* **19**(2), 213–229 (2016)
31. Prasopoulou, E.: A half-moon on my skin: a memoir on life with an activity tracker. *Eur. J. Inf. Syst.* **26**(3), 287–297 (2017)
32. Singha, D., Tripathi, G., Jara, A.J.: A survey of Internet-of-Things: future vision, architecture, challenges and services. In: *Proceedings on 2014 IEEE World Forum on Internet of Things (WF-IoT)* (2014)
33. Shin, D.-H.: Ubiquitous city: urban technologies, urban infrastructure and urban informatics. *J. Inf. Sci.* **35**(5), 515–526 (2009)
34. Sundmaeker, H., Guillemin, P., Friess, P., Woelffle, S.: Vision and challenges for realising the Internet of Things. *Eur. Comm.* **3**, 34–36 (2010)
35. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of Things for smart cities. *IEEE Internet Things J.* **1**(1), 22–32 (2014)
36. Zhang, J., Puron-Cid, G., Gil-Garcia, J.R.: Creating public value through open government: perspectives, experiences and applications. *Inf. Polity* **20**(2), 97–101 (2015)
37. Teo, T., Srivastava, S., Jiang, L.: Trust and electronic government success: an empirical study. *J. Manag. Inf. Syst.* **25**(3), 99–131 (2008)
38. Vicente, P., Lourdes, T., Royo, S.: Are ICTs improving transparency and accountability in the EU regional and local governments? An empirical study. *Publ. Adm.* **85**(2), 449–472 (2007)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Strategies for Reducing Power Consumption and Increasing Reliability in IoT

Ricardo Reis^(✉)

Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS),
Caixa Postal 15.064, 91501-970 Porto Alegre, RS, Brazil
reis@inf.ufrgs.br

Abstract. The Internet of Things (IoT) demands new challenges in the design of computing and electronics components. One of the challenges is the power reduction of this expanding network of connected devices, where the majority is permanently connected. In a large set of applications, another significant issue is reliability, especially on critical areas as health and transport. This paper shows an overview of design strategies that we have developed to reduce power consumption and to increase reliability in circuits that are components of the IoT, as the reduction of the number of transistors in IoT devices, using optimisation techniques and the physical design of circuits tolerant to radiation effects.

Keywords: Internet-of-Things · Optimization · Physical design · Fault tolerance · Radiation effects · Nanoelectronics

1 Introduction

The growing number of connected devices in the Internet of Things (IoT) is one of the reasons for the ever increasing increase in the number of transistors produced annually in the world. Figure 1, based on (SIA 2005), shows the number of transistors manufactured annually in the world, year by year. This impressive growth is due to 3 main factors: the increasing number of transistors integrated into a chip, the growing number of products that include embedded chips and the increasing number of manufactured copies of each product. The manufacturing cost of a transistor is relatively cheap. In (The Economist 2010) a comparison is presented between the cost of a grain of rice and the cost of a transistor. The price of a rice grain can be equivalent to the manufacturing cost of more than 125,000 transistors. This would indicate that there is no need to optimise the number of transistors in a design, since the cost of them is relatively small. But the cost of energy required for the operation of a transistor is increasing a lot. We also have to consider that a high-power consumption can reduce the lifetime of a system, as well as increase the effects of variability that can cause an integrated system to malfunction and/or also reduce its useful life. With the increasing connection of electronic and computational devices on the Internet, that is, in the Internet of Things, power consumption problems tend to get worse, and a lot. How much Power Plants we will need to cope with the IoT/IoE (Internet of Everything) world? This is a major issue.

So, an essential keyword on the Internet of Things is **optimisation**, especially the optimisation of power consumption, which must be addressed at all levels of abstraction in the design flow of a computer or electronic system. The total power optimisation is a summation of the optimisation done at each level of design abstraction. So, sustainable computing requires optimisation at all design levels of a computer or electronic system design.

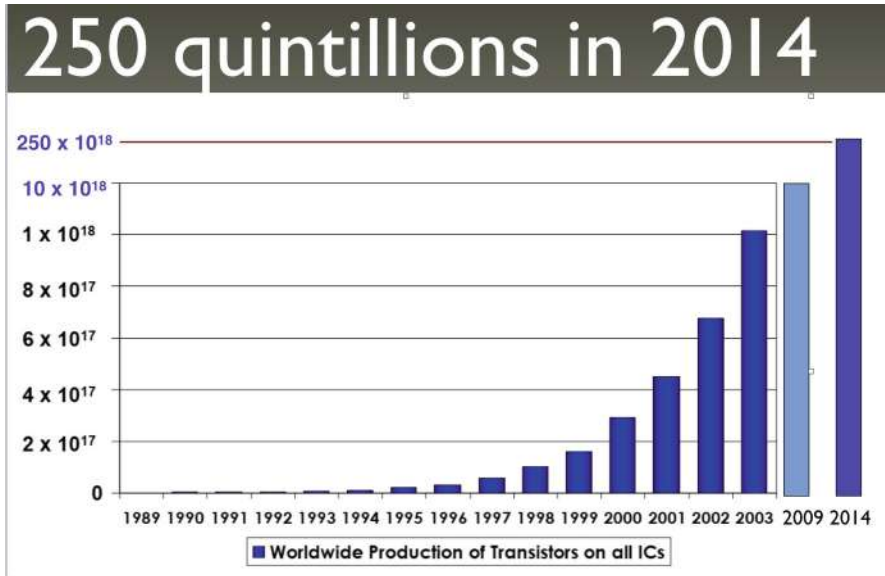


Fig. 1. Number of transistors produced annually in the world (adapted from SIA 2005)

2 Internet of Things

The term Internet of Things has already given rise to several other associated terms, such as the Internet of Health (IoH), Internet of People (IoP) and the Internet of Everything (IoE). In fact, the latter term becomes the most comprehensive, but each one of the others has some specific characteristics. When talking about the Internet of Health, which includes real-time monitoring of a person's clinical conditions, as well as chips injected in a person, the issue of reliability is a key one. And reliability is also related to power consumption in most cases. High power consumption can reduce the lifetime of a system. When it comes to the Internet of People, the issue of people's security and privacy is of great relevance. But in all cases, the importance of optimising energy consumption is growing more and more.

When considering optimisation, it means that integrated systems must increasingly be dedicated to the intended application to optimise the number of components, that means the number of transistors. Another important strategy for optimisation is the hardware and software codesign, where one can manage the compromise between performance, consumption, and reliability.

Devices connected to the Internet of Things (or the Internet of Everything), can have very different complexities. If it is analysed the complexity considering the number of components, we can find small devices with few transistors and large devices with billions of transistors. Of course, large devices will consume much more power, but we have to consider that most devices on the Internet of Things are devices with a low number of transistors. But, because they are found in large quantities, they can represent a total consumption more important than the consumption of the so-called large devices that are present in a lower number. Therefore, consumption optimisation must be performed on both large and small devices that are present in large quantities. Another aspect to consider is that some devices require the application of reliability techniques (such as those related to transport or health systems), which can increase the number of components, while other devices are not critical, such as a camera or video, where an error in viewing a pixel of an image does not cause significant problems.

Also, we can expect that many systems connected to the Internet of Everything will be Cyber Physical Systems (CPS), that are systems composed by different classes of components like electronic elements, mechanical elements, optical elements, physical sensors, chemical sensors, organic components, and many others. So, it is needed to obtain EDA tools to cope with the design of CPS composed of all these classes of devices.

Figure 2 (The Connectivist 2014) shows an estimate of the number of devices connected to the Internet since 1992 when they were about 1 million devices. By 2020 when it is estimated that there will be more than 50 billion devices connected in the network, and there are currently around 35 billion connected devices. In (IHS Markit 2018) the number of devices connected to the network in 2018 is shown by industrial and commercial sectors, where almost half is in the area of communication. The significant growth in the number of connected devices to the Internet has naturally led to a considerable increase in the energy consumed in the Internet of Things. For how long will we have the energy to meet this growing demand? Therefore, it is necessary to use techniques to minimise the energy consumption of each connected device in the Internet of Things.

The Internet of Health (IoH) is a significant way to increase the life of human beings but also to improve life quality. Some of the examples of devices to be connected to the IoH are: Glasses that can advise eye correction; Toothbrush that can find cavities and breath issues; Razor that identify acne; Pacemakers that broadcast data to cardiologist; Underwearables that can provide early detection of cancer and other anomalies; Combs that can scan for fungus and hair loss; Earphones that does measurement of hearing, analysis of emotional level; Watches able to measure parameters like blood pressure, heart rate and others.

In critical areas such as the design of implanted devices (chips) in humans (Fig. 3), the reliability of the implanted systems is obviously critical. Some of the techniques used are based on the triplication of circuits and the temporal analysis of the propagation of a signal. Previously, the design of fault-tolerant circuits, to cope with radiation effects, was mainly in circuits that were sent to space. With the reduction of the value of the supply voltage of integrated circuits, nowadays the integrated circuits for use at ground level are also sensitive to errors caused by the radiation incident on the earth. Therefore, in critical areas such as implanted chips in humans, it is necessary to implement radiation effects tolerance techniques (Velazco et al. 2007). Also, critical

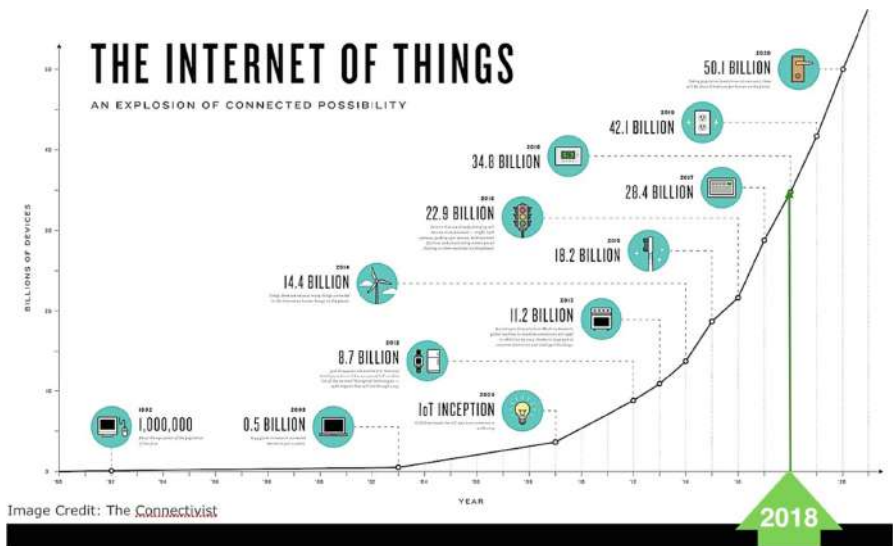


Fig. 2. Number of devices connected on the Internet (adapted from The Connectivist 2014)

systems used on the Internet of Health should be tolerant to any kind of noise (internal or external to the human body). They also must have a larger lifetime as possible, for obvious reasons and also should cope with environmental variability.



Fig. 3. The implantation of Chip Systems in humans demands reliability and ultra-low consumption

Also, there is the effect of “ageing”, that is, the ageing of the circuit, which is more eminent in nanometric technologies (Vasquez et al. 2012). One of the most important effects is known as NBTI (Negative Bias Temperature Instability) that alters the threshold voltage of the PMOS transistors, degrading the operation of the transistor. Another effect that causes failures in circuits throughout their life is the effect of electromigration, which can cause short circuits or rupture of connections (Fig. 4). In order to increase the lifetime of the chips, it is necessary to use physical design techniques that reduce the probability of electromigration (Posser et al. 2016, 2017).

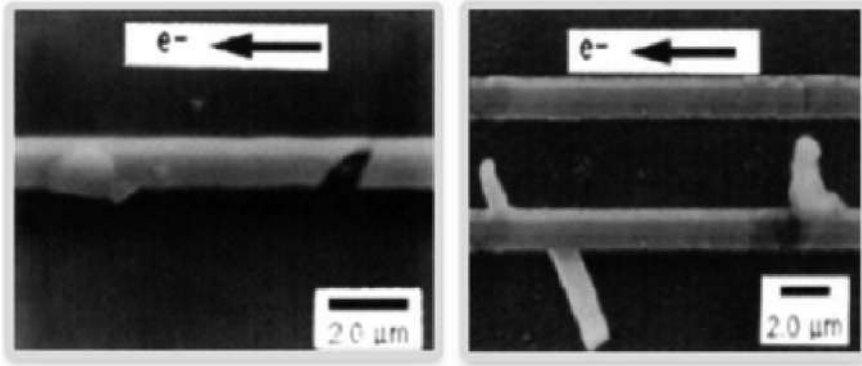


Fig. 4. Exemplo of a Void (open circuit) and hillock (short circuit) (Geden 2011).

3 Electronic Design Automation (EDA) Tools

The use of EDA tools is essential for optimising energy consumption and increasing reliability, as the design flow has a large set of steps as well the number of components of a chip can reach billions of transistors. In Fig. 5 we can see the floorplan of an integrated circuit, where the hotter colours show regions (hot spots) with higher energy consumption, indicating that in some points there is a significant concentration of power consumption. One way to deal with the problem is to modify the placement of the logic cells in the circuit to distribute the cells with the highest energy consumption over the entire circuit area. But this must be done without compromising the area, wirelength and operating frequency specifications (much depends on the routing). Another way is to decrease the number of transistors, since the static consumption is related to the number of transistors (Reis 2011A).

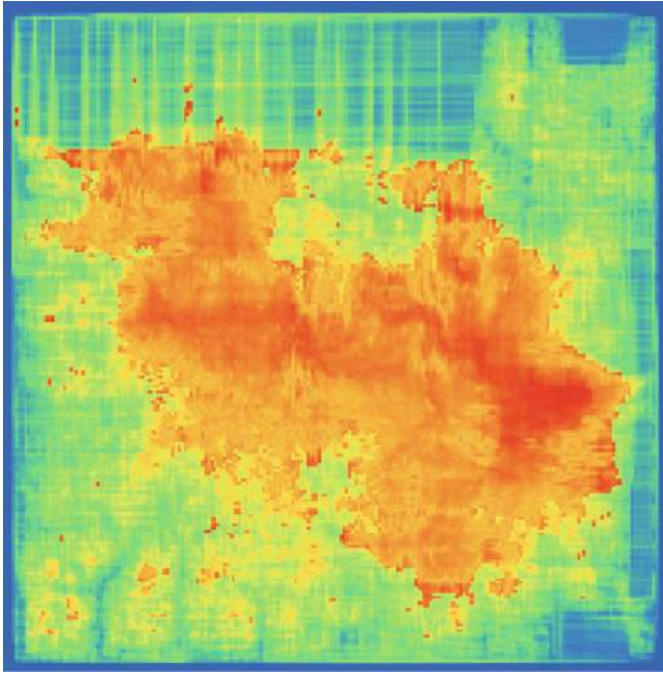


Fig. 5. View of the consumption density on a chip

4 Power Consumption Reduction by Reducing the Number of Transistors

The reduction of the power consumption of a System on a Chip (SoC) is a function of a sum of techniques and strategies of design applied in different levels of abstraction in the design flow of an integrated system (Reis 2010). The summation of the gains is that it will set the total gain in power reduction. When we deal with the physical synthesis of a system on a chip, one technique is the optimisation of the number of components, that is, the number of transistors. In Fig. 6 (Reis 2011A) we can observe two solutions for the implementation of the same equation. The first solution makes use of 4 basic logic gates (3 NOR 2-input ports and one CMOS inverter), using a total of 14 transistors. The second solution makes use of only one logic gate, which performs the same function but with only 8 transistors. That is, the second solution, having a reduction in the number of transistors, will also have a proportionally smaller static power consumption. Furthermore, in the example of Fig. 6, we can see that the first solution also has 3 connections between the basic gates (and therefore even vias and contacts) that are eliminated in the second option with only one logic gate.

This elimination of connections is increasingly important because it decreases the number of connections to be implemented using the different metal layers. The decrease in the number of connections decreases the density of connections and, therefore, increases the routability of the circuit and also contributes to reduce the

average length of the connections, which implies in a reduction of the delay. In modern technologies, the delay in connections is so or more significant than the delay in the switching of logic gates. A greater spacing between the connections also contributes to an increase of reliability, due, for example, to the reduction of the possibility of electromigration, as already mentioned above.

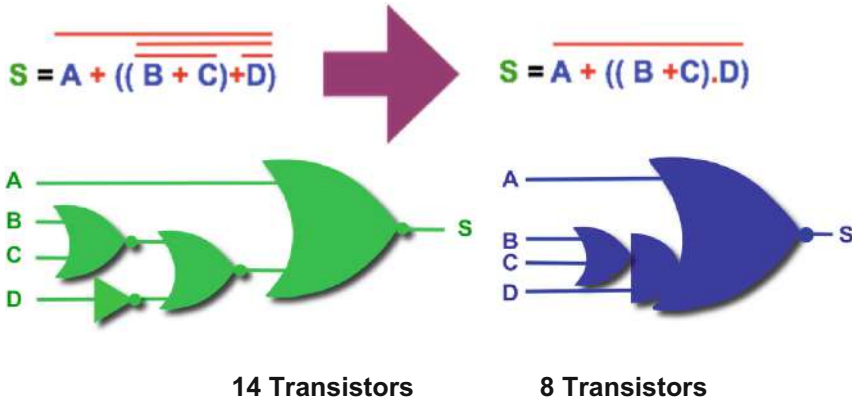


Fig. 6. Two options for the implementation of the same function (Reis 2011A)

The reduction of the number of transistors depends on the use of efficient Electronic Design Automation (EDA) tools that transform the logical equations of a system so that in addition to mapping equations in CMOS gates, make optimum use of complex logic gates. In (Conceição et al. 2017) we present a tool to reduce the number of transistors in a circuit through the fusion of networks of transistors that present fanout equal to 1. Also, it is fundamental the use of an automatic synthesis tool that can perform the automatic layout of any logical function. There is no use to achieve a logical optimisation if it is necessary to map (transform) the equations according to the logic gates available in a traditional cell library [which have few functions, in general, no more than 100 functions], as is still done when using traditional EDA systems. This mapping step is called technology mapping, and it represents a step of deoptimization. With this aim, we have developed automatic layout synthesis tools such as ASTRAN (Ziesemer and Reis 2015) (Fig. 7), which allows automatic generation of the layout of any network of transistors (Reis 2011A, B).

Another technique to reduce consumption is through the sizing of the transistors. Modern integrated circuit manufacturing technologies show a significant increase in static power consumption that is often greater than dynamic power consumption. One way to mitigate power consumption, especially the static one, is to carry out a sizing of transistors to optimise power consumption (Posser 2011). In (Reimann et al. 2016) significant decreases in consumption are obtained through the use of automatic

transistor sizing tools. This is also called cell selection, where the cells are selected from a cell library. In this case, cell selection means the selection of cells with a specific size and V_{th} (threshold voltage). In traditional cell libraries, one function has in general 3 sizings (one for less area, one for less power, and one for less delay) and 3 V_{th} (threshold voltage).

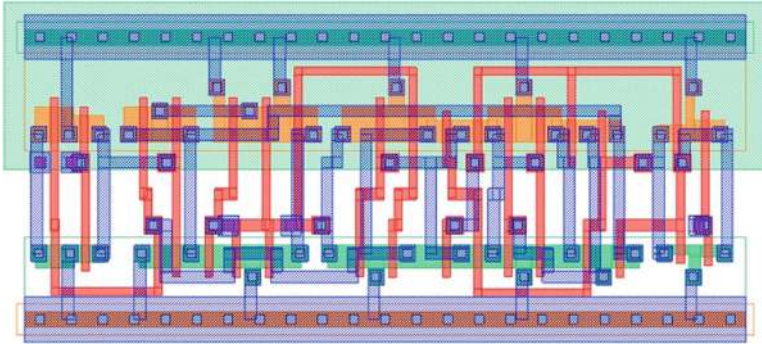


Fig. 7. Transistor network layout generated automatically (Ziesemer and Reis 2015)

5 Reliability

As in the reduction of power consumption, in the design of critical systems, it is needed to use techniques to increase reliability at different levels of design abstraction. At the architectural level, a very applied method is the redundancy of modules, especially triple module redundancy (TMR) (Kastensmidt et al. 2006). Another is the temporal redundancy (Nicolaidis 1999) where a signal traverses two paths, one with higher delay and another one with less delay. The difference of delay must be longer than the duration of a transient. Comparing the signal after traversing the two paths indicates whether there has been a transient propagation or not. At the physical level, we can apply different techniques to reduce or avoid problems such as electromigration (Posser et al. 2016, 2017). In the example of Fig. 8, the position of the output pin in the centre (point 4) increases the lifetime of the circuit because it allows reducing the maximum density of current in the segments of the metal layer.

In (Velazco et al. 2007) it is presented a series of works aimed at mitigating the effects of radiation on integrated circuits. In (Kastensmidt et al. 2006; Neuberger et al. 2014; Gennaro et al. 2017; Aguiar et al. 2016; Lazzari et al. 2011) we present some of the results that our research group has obtained in the development of techniques aiming the design tolerant to faults due to transients, as the effects due to radiation.

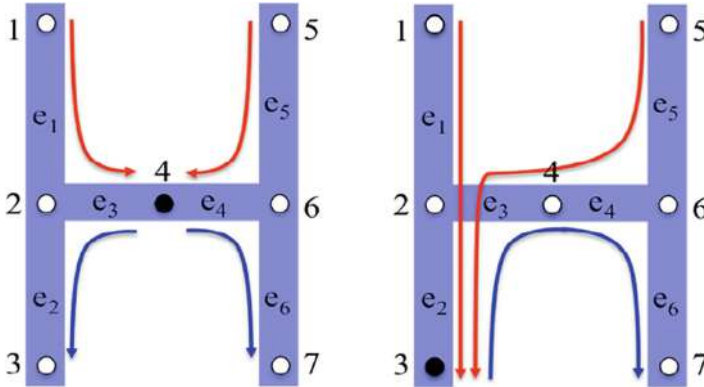


Fig. 8. Changing current density with the change of position of the output pin (Posser et al. 2016, 2017)

6 Hardware Accelerators

The evolution of computer architectures, that today means, the evolution of microprocessor architectures has been very significant. In the 1970s, one marketing argument from microprocessor producers was the number of instructions that the microprocessor could execute as well as the clock frequency of the microprocessor. In the last decades, there has been a change of paradigm, discontinuing the race for the increase of the clock frequency, because the increment of the clock means an increase of the dynamic consumption. Instead, there was an increase in the number of cores (CPUs) aiming at increasing performance. Initially with homogeneous cores and later with heterogeneous cores.

Currently, we can find chips with multiple CPUs and several GPUs (as can be seen in Fig. 9 (Shao 2016) showing the floorplan of the A8 microprocessor (from Apple). In this same figure, it can be observed that about half of the area is occupied with hardware accelerators, which are modules dedicated to the execution of a specific function. For example, an encryption module placed next to the output/input pins and which will encode the output data and decode the received data. So, the execution of this function will be faster, because it is done by a dedicated module (that means smaller) and with only the needed number of components to perform that function. It also will consume less power.

A more important fact is that the use of hardware accelerators leads to greater energy efficiency (allowing more sustainable computing), mainly due to the reduction in the number of components used to perform a function. At any given time, only the hardware accelerators in use at that time are being powered. So, the hardware accelerators that are not in use are disconnected from the power supply. This strategy is also known as “Dark Silicon”. We can even predict architectures consisting essentially of hardware accelerators, with only one or two small CPUs to manage these hardware accelerators.

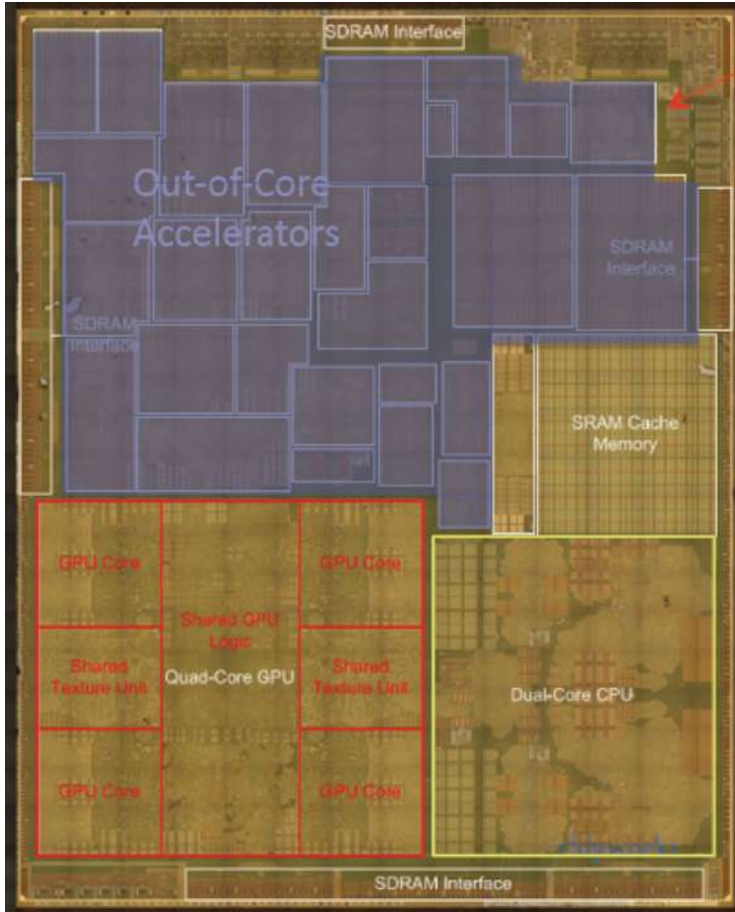


Fig. 9. Apple 8 floorplan with 29 hardware accelerators (AnandTech [2014](#); Shao [2016](#))

The introduction of an NPU in A11 is another element characterising the heterogeneity of the SoC (chip system). And we can expect increasingly heterogeneous architectures, with dedicated modules for different operations to be performed by a SoC. In Fig. 10 (Techinsights [2017](#)) the floorplan of the Apple A11 microprocessor is presented, where one of the modules is an NPU (Neural Processing Unit). NPU is mostly dedicated to facial recognition (Techinsights [2017](#)), processing machine learning tasks more efficiently, consuming less energy than CPUs do. The CPUs occupy about 15% of the area of the chip and 6 GPUs occupy about 20% of the area. Most of the area is filled with the hardware accelerators. That is, it is growing in the architecture of Apple microprocessors the use of hardware accelerators.

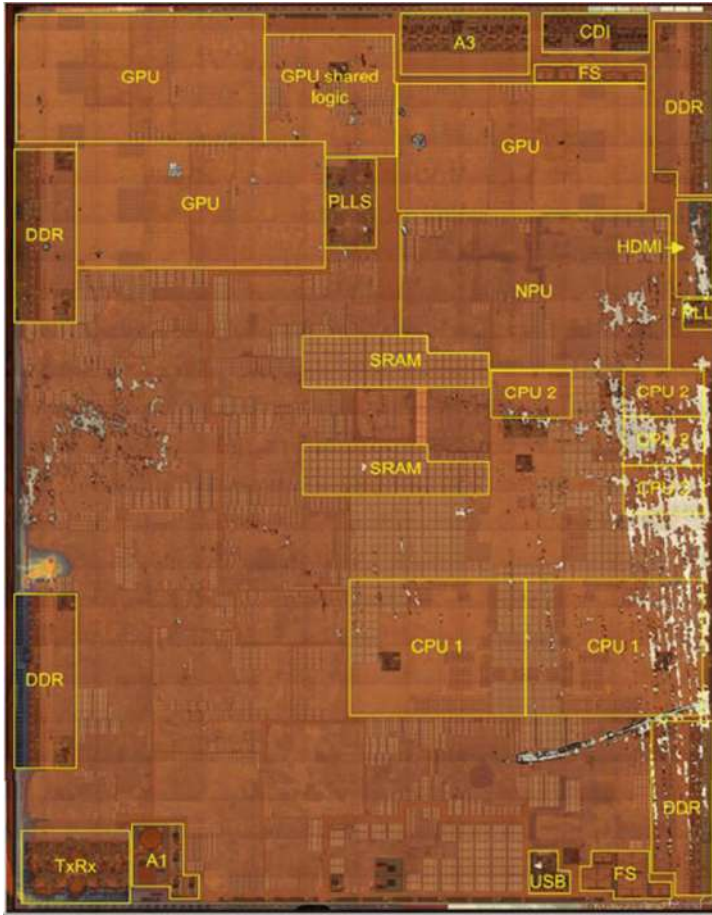


Fig. 10. Apple 11 floorplan with an NPU (Techinsights [2017](#))

7 Conclusions

To have sustainable computing, when the number of connected devices in the Internet of Things is fast increasing, it is fundamental the design of devices optimised regarding energy consumption. Most of the chips produced today use much more transistors than necessary to perform a function. So, there is a significant space for the optimisation of the number of components. In many devices related to critical applications, the application of techniques for fault tolerance is also fundamental, as nowadays circuits at ground level can have faults due to radiation effects. The reduction of power consumption must be treated at all design abstraction levels in a synthesis flow of integrated systems, from the specification of them in high-level languages to the physical synthesis. It was presented several works that were developed to reduce the power consumption and increase the reliability of integrated systems on a chip, and more

details are shown in the mentioned references. The keyword in the age of the Internet of Things is **optimisation**.

Acknowledgements. We thank CNPq, FINEP, Fapergs, and CAPES for financial support for the development of our team's work, as well as the master's and doctoral students of PGMICRO and PPGC and students of Scientific Initiation who have contributed to the research works that served as the basis for this paper.

References

- Aguiar, Y., Zimpeck, A., Meinhardt, C., Reis, R.: Permanent and single event transient faults reliability evaluation EDA Tool. In: *Microelectronics Reliability*, September 2016, vol. 64, pp. 63–67. Elsevier B.V., Amsterdam (2016). ISSN 0026-2714
- AnandTech (2014). <https://www.anandtech.com/show/8562/chipworks-a8>
- Conceição, C., Moura, G., Pisoni, F., Reis, R.: A cell clustering technique to reduce transistor count. In: *24th IEEE International Conference on Electronics, Circuits and Systems – ICECS 2017*, Batumi, Georgia, 5–8 December 2017, pp. 186–189 (2017). <https://doi.org/10.1109/icecs.2017.8291996>
- Gennaro, R., Rosa, F., Oliveira, A., Kastensmidt, F., Ost, L., Reis, R.: Analyzing the impact of fault tolerance methods in ARM processors under soft errors running Linux and parallelization APIs. *IEEE Trans. Nucl. Sci.* **64**(8) (2017). <https://doi.org/10.1109/tns.2017.2706519>. ISSN 1558–1578
- Geden, B.: Understand and avoid electromigration (EM) & IR-drop in custom IP blocks. Synopsys (2011)
- Lazzari, C., Wirth, G., Kastensmidt, F., Anghel, L., Reis, R.: Asymmetric transistor sizing targeting radiation-hardened circuits. *J. Electr. Eng.* (2011A). <https://doi.org/10.1007/s00202-011-0212-8>. Accessed June 2011
- Kastensmidt, F., Carro, L., Reis, R.: *Fault-Tolerance Techniques for SRAM-Based FPGA*, pp. 1–183. Springer, New York (2006). <https://doi.org/10.1007/978-0-387-31069-5>. ISBN 0-387-31068-1
- Neuberger, G., Wirth, G., Reis, R.: *Protecting Chips Against Hold Time Violations Due to Variability*, pp. 1–107. Springer, New York (2014). <https://doi.org/10.1007/978-94-007-2427-3>. ISBN 978-94-007-2426-6
- Nicolaidis, M.: Time redundancy based soft-error tolerance to rescue nanometer technologies. In: *Proceedings of IEEE VLSI Test Symposium*, vol. 17, pp. 86–94. IEEE Computer Society (1999)
- Posser, G., Flach, G., Wilke, G., Reis, R.: Gate sizing minimizing delay and area. In: *ISVLSI 2011, IEEE Computer Society Annual Symposium on VLSI*, Chennai, India, 4–6 July 2011, pp. 315–316 (2011). <https://doi.org/10.1109/isvlsi.2011.92>. ISBN 978-0-7695-4447-2
- Posser, G., Mishra, V., Jain, P., Reis, R., Sapatnekar, S.: Cell-internal electromigration: analysis and pin placement based optimization. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.* **35** (2), 220–231 (2016). <https://doi.org/10.1109/TCAD.2015.2456427>. ISSN: 0278-0070
- Posser, G., Sapatnekar, S., Reis, R.: *Electromigration Inside Logic Cells*, 118 p. Springer (2017). <https://doi.org/10.1007/978-3-319-48899-8>. ISBN: 978-3-319-48898-1
- Reimann, T., Sze, C., Reis, R.: Challenges of cell selection algorithms in industrial high performance microprocessor designs. In: *Integration*, vol. 52, pp. 347–354. Elsevier B.V., January 2016. <https://doi.org/10.1016/j.vlsi.2015.09.001>. ISSN 0167-9260

- Reis, R.: Redução de Consumo pela Otimização de Componentes. In: SEMISH 2010, Anais do 37º Seminário Integrado de Software e Hardware, Belo Horizonte, 21 a 22 de julho de 2010, pp. 371–379 (2010). ISSN 2175-2761
- Reis, R.: Design automation of transistor networks, a new challenge. In: IEEE International Symposium on Circuits and Systems, ISCAS 2011, Rio de Janeiro, Brasil, 15–19 May 2011, pp. 2485–2488. IEEE Press (2011A). <https://doi.org/10.1109/iscas.2011.5938108>. ISBN 978-1-4244-9472-9
- Reis, R.: Power consumption & reliability in NanoCMOS. In: IEEE NANO, 11th International Conference on Nanotechnology, Portland, USA, 15–19 August 2011 (invited talk), pp. 711–714 (2011B). <https://doi.org/10.1109/nano.2011.6144656>. ISBN 978-1-4577-1515-0
- The Connectivist (2014). <http://ow.ly/i/5vph6/original>
- The Economist (2010). Accessed 6 Sept 2010
- SIA: Semiconductor Industry Association, Rrebooting the IT Revolution (2015). <http://www.semiconductors.org/clientuploads/Resources/RITR%20WEB%20version%20FINAL.pdf>
- IHS Markit: IoT Trend Watch 2018 (2018). https://ihsmarkit.com/forms/thankyou.html?efid=t+m2jEyFYkYQYyoP3YvuHA==&&gasc_id=862037098&&gasc_label=scrXCLnM7m0Q6siGmwM
- Techinsights (2017). <http://techinsights.com/about-techinsights/overview/blog/apple-iphone-8-teardown/>
- Vazquez, J., et al.: Delay sensing for long-term variations and defects monitoring in safety-critical applications. *Analog Integr. Circ. Sig. Process.* **70**(2), 249–263 (2012). <https://doi.org/10.1007/s10470-011-9789-0>. ISSN 0925-1030
- Velazco, R., Fouillat, P., Reis, R.: Radiation Effects on Embedded Systems. Springer, New York (2007). ISBN 978-1-4020-5645-1
- Shao, Y.S.: Design and modeling of specialized architectures. Ph.D. thesis, Harvard, May 2016. <https://ysshao.github.io/papers/shao2016-dissertation.pdf>
- Ziesemer, A., Reis, R.: Physical design automation of transistors network. *Microelectron. Eng.* **148**, 122–128 (2015). <https://doi.org/10.1016/j.mee.2015.10.018>. ISSN 0167-9317

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





An Internet of Things (IoT) Model for Optimising Downtime Management: A Smart Lighting Case Study

Brenda Scholtz¹✉, Mando Kapeso¹, and Jean-Paul Van Belle²

¹ Nelson Mandela University, Port Elizabeth, South Africa
{brenda.scholtz, mando.kapeso}@mandela.ac.za

² University of Cape Town, Cape Town, South Africa
jean-paul.vanbelle@uct.ac.za

Abstract. In today's global, competitive economy, downtime has been identified as a key performance indicator for field service organisations. The emergence of an Internet of Things (IoT) has brought new enhancement possibilities to various industries such as the manufacturing and field service industry. This paper provides a vision and motivation for using IoT in Field Service Management (FSM) in order to address data quality and service delivery issues. The theory of information quality was used to undergird the research and a model for the optimisation of downtime management in the field service industry using the IoT is proposed. The model was used to drive the design of a "proof of concept" prototype, the KapCha prototype. The paper also includes a report on an empirical study of the application of the proposed IoT model in FSM. The experiment findings showed that the prototype reduced the round trip delay time for sending and receiving data and was scalable. As a result, access to quality information supporting advanced data analytics and artificial intelligence was provided. Therefore, service technicians can be alerted more quickly as soon as any potential technical problems occur. In turn improved diagnostics and more efficient decision making can be achieved. The model and the lessons learned provide valuable guidance to other researchers and fill the gap in research of empirical studies conducted on IoT implementations.

Keywords: IoT · Smart lightning · Design science research

1 Introduction

The lack of accurate, available and real-time information is a common challenge faced by field service organisations [1]. As a result downtime, which is an important performance measure in this domain, is negatively impacted. Downtime is defined as the time between a customer's request for service and the completion of the service by the field service team to rectify the problem [2]. Field service management (FSM) refers to the support provided by hardware and software in the management of field service operations and involves the management of the activities and processes that are associated with field services. There is a need for solutions that efficiently address the challenges of FSM and downtime management and that support the provision of

quality information and ultimately improved decision making and service delivery levels [3].

One application of downtime management in the field service domain is that of the maintenance of smart lights. Smart lighting projects have been undertaken by municipalities as a result of a drive for improved energy management within cities [4]. In the context of cities, streetlights are one of the most important assets to maintain as they provide safe roads and enhanced security for homes, businesses and city centres. However, they are costly to operate and account for an estimated 40% of the amount of electricity spent in an urban city [5]. To address this issue, city managers are implementing smart lighting solutions. Smart lighting consists of heterogeneous and multidisciplinary areas of lighting management, with the possibilities of integrating a wide range of sensory and control technologies with ICTs. This integration can improve efficiencies in lighting products and lower the negative impact derived from the use of energy for illumination. Smart lighting provides intelligent features and interfaces for lighting solutions in the ambient, commercial and public domains [4, 6]. Smart lighting is linked to the concept of a smart city, which is an urban development that envisions the efficient management of a cities resources and services with the use of integrated ICT solutions [4]. Smart cities play an important role in the sustainable economic development of countries or states seeking to attain environmental sustainability. Smart cities are made possible through the abundance of smart devices, smart objects and the emergence and rapid growth of technologies such as the Internet of Things (IoT). The IoT is described as a decentralised system of “smart” objects with sensing, processing, and network capabilities [7].

Extensive research has been conducted related to the IoT [16, 18, 26]. In particular several studies have proposed various architectures for IoT, such as a general reference architecture [27] and others in certain domains such as smart metering [28]. However, limited studies can be found that report on empirical studies of IoT applications in practice and findings and lessons learnt from these applications. There is a need for research into how technologies in the IoT can be applied to various business domains [20].

This paper addresses this gap by investigating an IoT application in the domain of smart lighting. The purpose of this paper is to propose an IoT model that addresses the challenges of information quality leading to poor downtime management. The paper reports on the application of this model in the smart lighting domain. The model includes IoT compatible technologies and techniques (protocols and formats) to support successful downtime management. To address this purpose, a critical analysis of the literature related to FSM, downtime management and IoT was conducted (Sect. 2). The context was a smart lighting organisation in South Africa (Sect. 3). From the literature and consideration of the context, a theoretical model was derived (Sect. 4). The model was used to design the architecture of and to implement a prototype for the case study (Sect. 5). The experiments conducted revealed that the new architecture and protocols implemented resulted in a lower Round Trip Delay time and was scalable (Sect. 6). The quality of information was improved and provided a foundation for advanced data analytics and artificial intelligence (AI), since the system provided intelligent information to technicians and managers; thereby improving diagnostic decision making, downtime management and service delivery.

There are several contributions and implications for future research that are identified from this study (Sect. 7). The practical contribution is the model, which can provide guidance to practitioners working in the field service domain and for system designers. On a theoretical level the model and the implementation issues identified contribute to the body of knowledge regarding the application of IoT models, architectures and network protocols.

2 Literature Review

2.1 Challenges in Field Service Management (FSM) and Downtime Management

In a competitive global economy where every organisation is looking at ways to cut costs, increase efficiency and gain a competitive advantage, organisations have become more customer-centric. The effectiveness of field services provided by technicians affects everything from the retention of customers and the profitability of the organisation [1, 8]. With field-based services, customers receive either an on-site or a remote service [2]. FSM operations include tracking vehicles, scheduling and dispatching employees, and integration of these operations with a back-office system for inventory, logistics and marketing. FSM includes elements such as Enterprise Asset Management (EAM), maintenance support, sensor networks, Radio Frequency Identification (RFID) tags, technical support, contract management and product life-cycle management. The FSM market has seen a steady growth and evolution in the last 10 years [9], which can be attributed to new technology developments, as technology is a driver in improved after-sales service innovations.

Downtime management is an important measure of performance for field services for both the organisation providing the service and the customer [10]. From the customer's perspective, that is the organisation undergoing downtime, the downtime period has operational implications such as reduced productivity levels and delayed delivery of services to the organisation's clientele. It is therefore imperative that downtime is kept to a minimal period. Service providers have to adequately manage downtime in order to satisfy its customers and by doing so efficiently they may gain a competitive advantage. Agnihothri [2] classified downtime into two subcategories, response time and on-site time. Response time is the time between the customer's request and the service team's arrival on-site. On-site time is the duration of time taken between the service team's arrival at the customer's site and the rectification of the problem. Corrective maintenance occurs when the machinery breaks down and includes activities undertaken to diagnose and rectify a fault so that the failed machine, equipment or system can be restored to its normal operational state, thus reducing the extent of downtime.

A lack of information related to a technical breakdown can result in longer cycle times and possibly a second service visit, thus resulting in longer periods of downtime for customers [8]. A malfunctioning piece of industrial machinery on a manufacturing floor can translate into tens of thousands of dollars per minute. It is important to make critical information immediately available to field technicians and management with

high levels of accuracy. Critical data related to the problem must be accurate, available anywhere, and dynamically changing along with the day-to-day operations of field service teams. Access to this information can assist with optimising the problem detection step in FSM and field service providers can determine strategies to ensure that downtime is minimised and managed with optimal efficiency. Within IS literature, information quality (IQ) can be used as a dimension of IS success [12]. Knowledge is functionally related to data and information, thus it follows a hierarchy (data \rightarrow information \rightarrow knowledge), termed as the knowledge hierarchy [11]. Our study classified the problems in FSM related to information that impact downtime management according to six of the attributes of IQ proposed by [12]. These are:

- Timeliness: lack of access to real-time information [1];
- Completeness: missing information [1, 3, 8];
- Accuracy: inaccurate information [1];
- Relevance: aggregated or de-aggregated information [14, 15];
- Consistency: lack of integration between enterprise and FSM systems [14].

This analysis also confirmed the findings of [13] showing a significant relationship between IQ and individual impact. Individual impact is measured in terms of decision-making performance, job effectiveness, and quality of work. Challenges faced by FSM organisations with regards to IQ resulted in negative impact on decision making and service delivery. Inaccurate or missing information and a lack of real-time availability of information to employees onsite in the field (for example dispatchers and service technicians) resulted in operational challenges [1, 3, 8]. Information related to the customer or the equipment under maintenance or repair is not always readily available to field service employees, resulting in the poor scheduling of field employees, the ineffective management of field service resources such as service parts [3] and ultimately in poor service delivery. In a study by Lehtonen [1], it was reported that service teams could not provide a service due to missing spare parts. The main reason for this was the inaccurate information on the spare parts that was taken to the client at the time of repair. Challenges in FSM within Enterprise Systems may also arise due to the lack of accessibility and integration of various systems [14]. For example, geographical data is found in Geographical Information Systems (GIS), whilst maintenance-related data and reports are often stored in an Enterprise Resource Planning (ERP) system, thus resulting in integration and consistency issues. Schneider [14] reported issues related to the use of aggregated data within an ERP system. For example, in an ERP system electricity usage data for a manufacturing plant is usually stored as an aggregated figure for all work centres within the plant. Aggregated data makes the operational performance monitoring of a single work centre or equipment within the plant difficult.

Access to real-time information aids organisations in optimising FSM since it can minimise the time for the service team to locate a client location by using GPS services and can reduce the on-site time spent servicing a clients' request [1, 2]. Real-time access to the clients' location eliminates the need for the service team to return to the service provider's facilities in order to get information about a new client's request, thereby optimising the scheduling element [3].

2.2 Applications of the IoT

The IoT has brought new functionality possibilities for many industries such as manufacturing and field services [16, 17]. It is expected that soon more than fifty billion devices ranging from smart phones, laptops, sensors and game consoles will be connected to the Internet through heterogeneous access network technologies [18]. However, the successful implementation of an IoT system introduces several other challenges. The abundance of data provided by sensors can introduce inefficiencies in data transfer and a need for aggregated data since sensor nodes are constrained by limited resources, for example computational power, memory, storage, communication, and battery energy [15]. These constraints provide an important challenge to design and develop approaches to information processing and aggregation that are efficient and make effective use of the data. For a given query, it may not be necessary or efficient to return all the raw data collected from every sensor – alternatively information should be processed and aggregated within the network and only processed and aggregated information returned. From a system level perspective, the IoT can be viewed as a dynamic, radically distributed, networked system, consisting of many smart objects that produce and consume information [19]. It can optimise business processes by leveraging on advanced analytics techniques applied to IoT data streams [19]. Thus, it provides good potential for addressing the downtime problem, if successfully implemented.

Although technology advances enable the possibility of the IoT, it is the application of the IoT which is driving its evolution [18]. The potential social, environmental and economic impact that the IoT has on the decisions we make and the actions we take is its main driving force. For example, having accurate information about the status, location and identity of things which are part of our environment opens the way for making smarter decisions. The application domains that the IoT includes can provide a competitive advantage beyond current solutions. In its inception the IoT was used in the context of supply chain management with RFID tags as the enabling technology [7]. However, in the past decade its applications have covered a wide range of industries, including transportation and utilities, to name just a few. Hwang et al. [20] classified the potential business contexts of IoT into three different factors: industry applications (for example government, education and finance); service domains (for example transportation, asset management) and value chain activities (for example sales and marketing, service or procurement). On the other hand, Borgia et al. [18], classified the IoT into three application areas: industrial (for example agriculture, logistics or other industrial applications), health/well-being and smart city. The smart city factor includes safety, mobility, buildings, road conditions, waste collection and public lighting.

3 Context of Research: Smart Lighting

The case study used in this research is a smart lighting system that is maintained at an engineering consulting and research organisation in South Africa. For purposes of anonymity, the organisation will be called LightCo. The smart lights that are used as

outdoor luminous equipment for parking bays and security lights for building facilities and are grid independent; meaning they are not connected to a local or municipal electricity provider for the energy needed to light them. An interview was conducted with one of the senior engineers at LightCo in order to establish an overview of the environment as well as the challenges faced by the organisation in delivering maintenance services for the smart lighting environment.

Smart lighting consists of the integration of intelligent functionalities and interfaces at four complementary levels [4], namely: the embedded level; system level; grid level and communication and sensing level. The embedded level is the lighting engine or the light itself, whilst the system level is the luminaries and lighting systems. The grid level consists of the management and monitoring of the power sources, energy generation and plants and the distribution of utilities and appliances. The final level is the communication and sensing level, which provides complete lighting solutions with monitoring, control and management of the applications.

The smart light unit at LightCo consists of an on board 48-voltage battery pack that is used as an energy storage unit. The solar panel is used to harness solar energy and the wind turbine generates electricity by the turning of a generator. The architecture of the smart lighting system allows for remote monitoring. The smart light also contains sensors and actuators that enable it to measure environmental variables and to respond to specific conditions by means of the actuators. The sensors include ambient sensors on the solar panel and voltage and current sensors on the circuit board of the smart light. Furthermore, the smart light is uniquely identifiable and contains on-board microcontrollers that provide computational and communication capabilities. The microcontrollers receive voltage and current data readings from the solar panel and wind turbine and also record the voltage and current that is outputted to the LED light. The battery management system manages the flow of current to the battery. Once these readings have been recorded they are then sent to a remote server for processing.

Prior to starting this study, the smart lighting system at LightCo did not provide for efficient or effective downtime management. Technical problems with the lights were not being reported timeously and were not correctly diagnosed due to IQ issues reported in literature [1, 14]. These problems could be for example, an LED light or circuit board is damaged. The system that was in place for detecting technical problems with their lights used a Global System for Mobile Communications (GSM) SMS-based messaging/polling protocol to transfer data from a smart light to a server at a remote location. This protocol was reported as inefficient due to its high latency times and high data costs affiliated with the sending and receiving of SMS messages. Increasing the latency was not an option, since it would increase the data costs. Data transmission was not bi-directional and data was merely recorded in a CSV file, with no processing performed on it. An Arduino microcontroller was situated in each smart light with a GSM Shield, which allowed the Arduino board to send and receive an SMS as well as connect to the Internet using the GSM library. However, the system did not use the GPRS wireless component that would enable the Arduino to connect to the Internet. Technicians had to manually peruse the data to diagnose any issues or potential issues.

4 IoT Model for Downtime Management

The Three Phase Data Flow Process model proposed by Borgia et al. [18] (Fig. 1), the four layers of IoT [25], and IQ theory were used as the main guiding theories for the proposed IoT Model for Downtime Management (Fig. 2). The model describes the flow of data in the IoT over three phases [18], namely the Collection Phase; the Transmission Phase; and the Process Management and Utilisation phase and four layers [25] (the Sensing Layer; the Networking Layer; the Service layer; and the Interface layer). The Sensing Layer consists of hardware that senses and controls the physical world and acquires data. Examples are RFID, sensors and actuators. The Network Layer provides networking support and transfers data over either a wireless or a wired network. The Service Layer is responsible for the provision of services to satisfy the user needs and creates and manages services. The Interface Layer (or Application Layer) interacts with other applications and users.

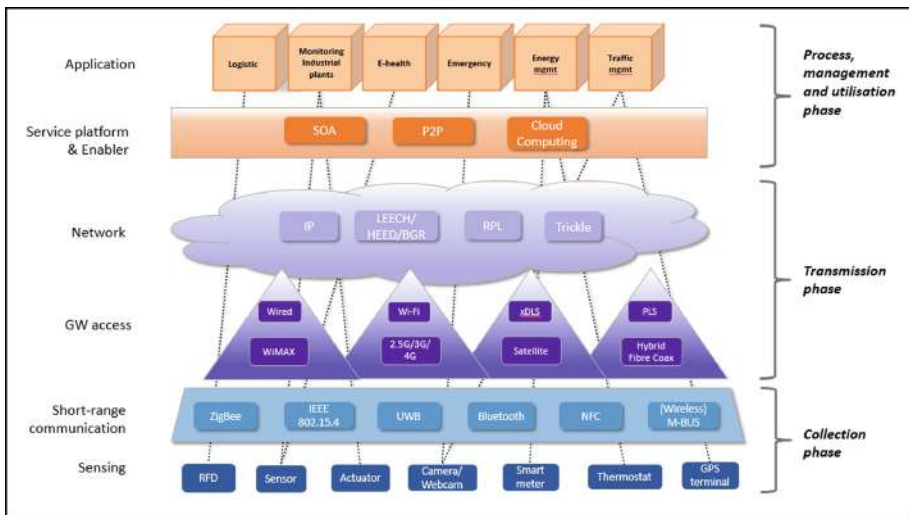


Fig. 1. The three phase data flow process [18]

The Collection Phase reports on the event driven processes during the collection and acquisition of the data from the environment [18]. Data acquisition technologies attached to sensors and cameras collect information about the physical environment (temperature, humidity and brightness), or about the objects (identify and state) in real-time; while data collection is accomplished by short range communications, which could be open source standard solutions or proprietary solutions. In the FSM context these would be integrated into the equipment or assets in the field, for example the smart light. The Transmission Phase involves mechanisms that deliver collected data to various applications and external servers [18]. Once data has been collected it must be transmitted across the network so that it can be consumed by applications. For wired

technologies the standard is Ethernet IEEE802.3. The primary advantage that wired networks have for data transmission is that they are robust and less vulnerable to errors and interference. However, they are costly. Therefore Wireless LAN (WLANs) are often used to access the network. Due to the flexibility of WLANs, it is believed that they will be the main communication paradigm of the IoT. However, the restricted wireless spectrum available for cellular networks is a major limitation to their wide-spread use.

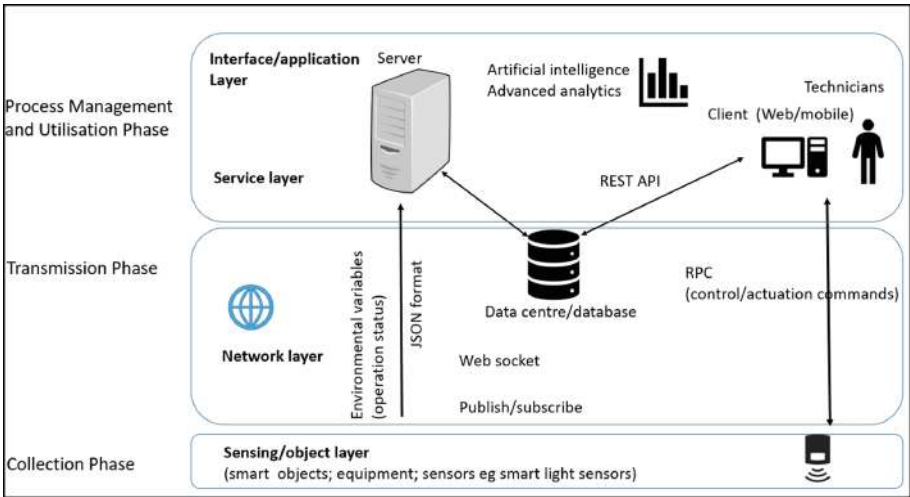


Fig. 2. IoT model for downtime management

The Processing, Management and Utilisation Phase incorporates the processing and analysing of information flows, data forwarding to services and applications and the provision of feedback to control applications [18]. It also involves device discovery and management, data filtering, aggregation and information utilisation. The Service Platform & Enabler sub-phase covers an important role for managing these functions and is necessary in order to hide the heterogeneity of hardware, software, data formats, technologies and communication protocols that are a key feature of the IoT. Its responsibility is to abstract all the features of objects, networks and services, and to provide a loose coupling of components.

5 Methodology and Development of the Prototypes

5.1 Methodology

The Design Science Research (DSR) methodology [21] was adopted in this study to create and evaluate the artefacts (model and prototype). The model was derived from a systematic literature review as well as from the case study of smart lighting

maintenance, which was used for implementing and evaluating the model. The Technical and Risk efficacy evaluation strategy from the Framework for Evaluation of Design Science (FEDS) is used in the DSR methodology for evaluations conducted in the design cycle of DSR and was used in this study to evaluate both the model and the KapCha prototype [22]. An artificial-summative evaluation was used to evaluate the design of the model, but due to space limitations these results are not reported on in this paper but are available on request. Iterative formative evaluations were conducted during the development of the prototype; after which a summative-naturalistic evaluation was conducted in order to determine the performance of the prototype under real-world conditions.

5.2 The Prototypes and Their Mapping to Requirements

The KapCha prototype was developed using an incremental prototyping process comprising of three prototype components (Table 1). **ProWebSoc** is the web socket protocol; **ProObjWeb** is the web socket client; and **ProDT** is the interface layer and web socket server. The IoT Model for Downtime Management (Fig. 2) was used to design the architecture of the prototypes.

Table 1. Prototype components

Collection and transmission phases		Process management and utilisation phase
ProWebSoc web socket protocol	ProObjWeb web socket client	ProDT (interface layer) web socket server
Data transmission protocols; security and standardisation	Transmission of data from the smart light and integration of CPS principles	Intelligent algorithms, advanced analytics and interfacing with applications and mobile technologies

Collection and Transmission Phases (ProWebSoc and ProObjWeb)

As an alternative to the SOAP/XML data transmission protocols used by LightCo prior to this intervention, a protocol based on JavaScript Object Notation (JSON) was implemented. JSON, is a text-based open standard format that is designed for human-readable data interchange and used for the serialisation of structured data making it easy for machines to parse and generate it. JSON is ideal for low processing computational capabilities (such as the smart light) and can result in less data that needs to be generated as compared to SOAP/XML.

ProObjWeb, through the web-socket client, enabled the smart light in the case study, as an OEM, to interface with a remote web-server using the KapCha web-socket protocol and to transmit data over a web socket protocol (**ProWebSoc**). Web sockets enable bi-directional communication (upstream and downstream) through the introduction of an interface and the definition of a full-duplex single communication channel that operates through a single socket [23]. They provide a reduction in network

traffic and latency as compared to polling and long-polling solutions that are used to simulate a full-duplex connection by maintaining two connections. They also reduce the amount of port openings on the server side, as compared to the traditional means of retrieving resources such as polling. This reduction also reduces maintenance of connection channels from the server side, therefore decreasing the overhead network traffic. The web socket protocol also has the ability to traverse firewalls and proxies, which is a problem for other protocols. The protocols provide real-time communication (RTC) between a smart object and a central system or other smart objects and supports ad hoc and continuous data transfer as well as operational status communication and Remote Procedure Calls (RPCs).

A GPRS wireless component was used to enable the Arduino to make use of web socket technology. The web socket client was developed on the Arduino board using the Arduino open source software and several web socket methods. During connection, the web socket detects the presence of a proxy server and automatically establishes a tunnel to pass through the proxy. The tunnel is established through the opening of a TCP/IP connection. The connection is established by the client issuing an HTTP connect statement to the proxy server for a specific host and port. Upon the tunnel being set up communication flows uninterrupted through the proxy.

The web socket protocol (**ProWebSoc**) was designed to work with existing web infrastructure, therefore the protocol specification defines that the web socket connection starts as an HTTP connection [24]. This guarantees full backwards compatibility with HTTP based communication protocols. The upgrade from an HTTP to web socket is referred to as a handshake. In this process the client sends a request to the server indicating that it wants to switch protocols from HTTP to Web sockets, by means of an upgrade header. During the handshake process the server accepts the request and responds with an upgrade switch header. The server acknowledges receipt of the client's request by taking the |Sec-Web socket-Key| value and concatenating it with a Globally Unique Identifier (GUID) in string form. An SHA-1 hash (160 bits) base64-encoded of this concatenation is then returned in the server's response. This prevents an attacker from tricking a web socket server by sending it carefully crafted packets using XMLHttpRequest or a form submission.

Web sockets are ideal due to the ability to use customised protocol calling depending on the service being offered [23]. In ProObjWeb, when the client receives a response with no errors the connection is upgraded to a web socket over the same TCP/IP connection. Once the connection is established data frames between clients and servers can be transferred. Once the web socket client application connects to the webserver, the webserver initiates an upgrade sequence to upgrade the connection from an HTTP connection to a web socket server. ProObjWeb was functionally tested using the web socket.org echo server, which allows developers of web socket applications to test the ability of their applications to successfully upgrade the connection from HTTP to a web socket protocol. The test results showed that ProObjWeb successfully managed to connect to the web.org server and upgrade from HTTP to web sockets.

Management and Utilisation Phase

The third prototype (**ProDT**) focused on the development of the web-socket server application, a decision tree algorithm implementation and a REST (Representational

State Transfer) API web interface. REST APIs with web socket Requests/Responses were used to form an intermediate layer between a client and the database, translate the raw data from the database to a format that the client requests and transmit the data. Most databases provide real-time notifications for added or updated data. However, real-time notification passage between the database and the client was required, therefore an Interface Layer was created in ProDT using a web socket server. The web-socket server also handled communication with the database, which is the core of the application architecture. These techniques (including the protocols) provided real-time notification between the database and the client, eliminating the need for the Ajax techniques of polling. The database contained information about each smart light such as the date of manufacture and installation, its GPS location, object data sent by the Smart Light, the fault issue result after data analysis and job card data.

In order to generate an issue, data analysis using a decision tree learning algorithm was implemented. A decision tree is a tree structure consisting of nodes that each represent a test of an attribute with each branch representing a result of the test [29]. The tree splits observations into mutually exclusive subgroups until observations can no longer be split. The ID4.5 is a popular splitting algorithm that builds a decision tree by employing a top-down, greedy search through the given sets of training data to test each attribute at every node. Decision trees require little effort for data preparation unlike some statistical techniques and they are easy to interpret. The data collected was categorised data and therefore is ideal for decision tree. Furthermore, as there was no historical information on the diagnosis of faults or issues the decision tree was the ideal AI technique to be used as the classifier would be developed from the expert's opinion of plausible issues. The diagnosis of a set of problems based on these opinions was determined as a classification problem.

The improvements in the quality of information provided by these techniques thus allowed for advanced data analytics and intelligent algorithms (such as decision trees) to be conducted on the IoT data streams. The ability to interface with mobile technologies was also provided.

6 Experiment Procedure and Findings

The aim of the experiments was to evaluate the Round-Trip Delay time (RTD) of messages, latency, accuracy of the decision tree analysis and scalability. Due to space constraints this paper only provides details of the RTD and accuracy experiments.

Round-Trip Delay time (RTD). The RTD experiment measures the time taken for a client to send a signal to a server and the time it takes for the server to acknowledge the signal and send a response [22]. In this context the client was the smart light and the server was the remote web socket server application. For the RTD experiment a connection had to be established between the smart light and the web socket server application. The experiment procedure involved running the applications in three cycles; the first cycle involved sending a data packet 10 times, then the second cycle 100 times and the third cycle 1000 times. The data packet was an array data object that was instantiated and sent to the server.

Two phases of experiments were performed: the first phase (local testing) consisted of running the server applications on a local host machine; and the second phase (remote testing) of the experiment involved running the server applications on a remote server. The performance metrics for the RTD evaluation were delay time and messages per second.

From the results of all three cycles it is evident that the KapCha web socket had a lower RTD time as compared to the Ajax protocol (Table 2). The results can be attributed to the fact that there are fewer HTTP overheads when using web sockets as compared to Ajax requests. Upon the connection being established all messages are sent over the single socket connection rather than the creation of new connections for new HTTP request and response calls created every time a message is sent over the Ajax protocol.

Furthermore, the web socket protocol had more messages sent per second as compared to that of the Ajax protocol. The messages per second for web sockets are higher because the web sockets establish the connection once over a single socket, unlike Ajax techniques that require multiple connections to be opened and closed during request/response calls. Therefore, web sockets do not have messages delayed during the connection process and can send more messages per second. The messages sent per second over the web socket protocol increased exponentially with the number of iterations completed.

Table 2. Experiment results – RDT and messages per second

First phase: local host					
Cycle	Num packets	Time (sec)		Messages	
		Ajax	Kapcha	Ajax	Kapcha
1	10	0.130	0.082	80.135	150.235
2	100	0.435	0.082	301.720	1502.235
3	1000	3.104	0.558	340.832	2215.621
Second phase: remote testing					
Cycle	Num packets	Time (sec)		Messages	
		Ajax	Kapcha	Ajax	Kapcha
1	10	1.241	3.705	3.212	1.752
2	100	22.016	17.213	2.907	3.352
3	1000	153.352	147.25	3.191	3.714

The results from the remote testing set of experiments revealed that the web socket protocol had a lower RTD time as compared to the Ajax protocol when the number of packets was lower than a certain level. This result could be due to the upgrade sequence overhead during the web socket handshake process. The additional overhead connection, however, is not significant as the number of iterations increase due to the maintenance of the single socket connection. The RDT results highlighted the advantages of applications that use web sockets have over HTTP polling mechanisms.

The advantages are lower latency and the provision of a single socket connection that enables the web server to push data to the client at will, creating a fully duplex bi-directional data exchange web-protocol.

Accuracy of decision tree: Prior to the development of the prototype there wasn't any data stored regarding the cause of a fault or the documentation of the diagnosis of a fault at LightCo. Therefore, the accuracy of the training dataset created was based on the expert's verification. The C4.5 decision tree algorithm was used to analyse the data and deduce the cause of the faults that occurred. For the experiment, the algorithm was deployed/executed to analyse three sample data set sizes of 50, 100 and 175. The number of correct predictions after each execution was recorded and verified by an expert at LightCo. This process was undertaken to establish the accuracy of the algorithm in diagnosing faults. The execution time of the algorithm was also recorded to determine the turnaround time of the fault diagnosis. The formula used for determining the accuracy percentage was:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{f_{11} + f_{100}}{f_{11} + f_{10} + f_{01} + f_{00}}. \quad (1)$$

The accuracy results are summarised in Table 3. The sample size of fifty ($n = 50$) resulted in a percentage accuracy of 82.2% meaning that 41 out of the 50 predictions were correct. The sample size of 100 had 79 correctly predicted faults with a percentage accuracy of 79%. The final sample size had an accuracy percentage of 77% that is 96 faults were correctly predicted. Whilst the accuracy results were all above 70% additional testing is required to determine accuracy of larger datasets. However, this could not be done since previous records were non-existent and the training set was small. This is a limitation of this study. Future studies should perform the accuracy tests on the larger data set, which will increase rapidly with time. However, in spite of this limitation useful results and lessons learned were obtained regarding the IoT techniques used in the model.

Table 3. Accuracy results

Size of dataset	Accuracy (%)
50	82
100	79
125	77
Mean	79

7 Conclusions

In this paper a theoretical prescriptive model for optimising downtime management is proposed that was derived from a systematic literature review of FSM, IoT and IQ theory. The use of intelligent algorithms and data accessibility are features of the model that can aid in the reduction of downtime. The model also supports geographically

dispersed devices and clients. From a practical viewpoint, an organisation in the smart lighting industry was used to test the model as a proof of concept. In the case of the smart lighting scenario, prior to the intervention of our study, an SMS/Ajax polling system was used that was slow and expensive due to the data costs. As a result, insufficient data was provided to assist with detecting and diagnosing problems. The solution lacked real-time information and field service technicians had to rely on human ‘diagnostics’ and sometimes travelling to the smart lights in order to physically detect problems. The proposed IoT model for downtime management was used to design an architecture and to develop and implement a system prototype for optimising downtime management in the smart lighting environment.

The evaluations of the prototype revealed that web-sockets are more efficient and cost-effective than other web-based data transfer protocols such as Ajax. The implementation of a web-socket based protocol provided a low-cost data communication protocol with real-time bi-directional capabilities and fully duplex communication between a smart light and a remote server. The use of IoT-enabled communication protocols reduced the latency time and data exchange costs. Furthermore, the web-socket server implements an expert system mechanism using intelligent algorithms for data analysis. The intelligent algorithm, a C4.5 decision tree, automates fault detection and provides an issue report. The intelligent algorithms can assist service technicians to identify and diagnose problems. The practical contributions of this research are therefore the model, which can be used by FSM organisations in the implementation of IoT. The results of the evaluations revealed that the implementation of the various techniques and features of the model optimised downtime within the smart lighting environment. A problem encountered during the study related to restrictions on GSM protocols by the mobile service providers, some of which do not support the use of web-socket connections. Another challenge was inventor patents on the smart lights in the case study that restricted testing of the prototype in its natural environment. As a result, only historical data was used for testing. A further limitation was that not all elements of the model could be tested due to time and resource constraints. However, the findings of this study can still be used by other researchers as a valuable source of reference when conducting similar research. The lessons learnt can be useful to other researchers and practitioners working in FSM and similar industries that can benefit from IoT.

The combination of advanced big data analytics, cloud-computing and IoT enables users to not only gather vast amounts of data but also enable them to process it without having to acquire high infrastructural costs. This leads to several opportunities for researchers in these fields. Future research directions could extend the study to include functionality such as predictive maintenance. AI mechanisms can be implemented in the model to support the prediction of faults before they occur. Additional intelligence can be achieved by interacting with other systems in the same environment that have a direct impact on the equipment’s performance. The addition of predictive mechanisms as well as enabling object interaction with other systems will transform regular equipment into a self-aware and self-learning machines, and consequently improves overall performance and maintenance management. The model serves as a reference model for standards and protocols in an IoT-based implementation in the field of downtime management within the after-sales industry. Although the study was limited

to evaluating the prototype in only one environment it provided valuable lessons that could be used by other practitioners and researchers to guide the implementation of IoT in FSM.

Acknowledgements. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF.

References

1. Lehtonen, O.: Taking advantage of after-sales product information in a multi-company environment. Masters thesis, Helsinki University of Technology, Department of Industrial Engineering and Management, Finland (2005)
2. Agnihothri, S., Sivasubramaniam, N., Simmons, D.: Leveraging technology to improve field service. *Int. J. Serv. Ind. Manag.* **13**(1), 47–68 (2002). <https://doi.org/10.1108/09564230210421155>
3. Petrakis, I., Hass, C., Bichler, M.: On the impact of real-time information on field service scheduling. *Decis. Support Syst.* **53**(2), 282–293 (2012). <https://doi.org/10.1016/j.dss.2012.01.013>
4. Castro, M., Jara, A.J., Skarmeta, A.F.G.: Smart lighting solutions for smart cities. In: *Proceedings - 27th International Conference on Advanced Information Networking and Applications Workshops, WAINA*, pp. 1374–1379 (2013)
5. Basu, C., et al.: Sensor-based predictive modeling for smart lighting in grid-integrated buildings. *IEEE Sens. J.* **14**(12), 4216–4229 (2014). <https://doi.org/10.1109/JSEN.2014.2352331>
6. Koh, L.H., Tan, Y.K., Wang, Z.Z., Tseng, K.J.: An energy-efficient low voltage DC grid powered smart LED lighting system. In: *IECON Proceedings (Industrial Electronics Conference)*, pp. 2883–2888 (2011)
7. Kortuem, G., Kawsar, F., Fitton, D., Sundramoorthy, V.: Smart objects as building blocks for the Internet of things. *IEEE Internet Comput.* **14**(1), 44–51 (2010). <https://doi.org/10.1109/MIC.2009.143>
8. Jose, G.J., Kumanan, S., Venkatesan, S.P.: Optimize field service management through analytics. In: *Proceedings of the International Conference on Advances in Production and Industrial Engineering 2015*, pp. 529–534 (2015)
9. Gartner. Magic Quadrant for Field Service Management (2017). <https://www.gartner.com/doc/3808464/magic-quadrant-field-service-management>
10. Knotts, R.M.H.: Civil aircraft maintenance and support. *J. Qual. Maint. Eng.* **5**(4), 335–348 (1999). <https://doi.org/10.1108/13552519910298091>
11. Nonaka, I.: A dynamic theory of organizational knowledge creation. *Organ. Sci.* **5**, 14–37 (1994)
12. DeLone, W.H., Mclean, E.R.: The DeLone and McLean model of information systems success: a ten year update. *J. Manage. Inf. Syst.* **19**, 9–30 (2003)
13. DeLone, W.H., McLean, E.R.: Information systems success: the quest for the dependent variable. *Inf. Syst. Res.* **3**(1), 60–95 (1992)
14. Schneider, J., et al.: Asset management techniques. *Int. J. Electr. Power Energy Syst.* **28**(9 SPEC. ISS), 643–654 (2006). <https://doi.org/10.1016/j.ijepes.2006.03.007>
15. Przydatek, B., Song, D., Perrig, A.: SIA: secure information aggregation in sensor networks. In *ACM SenSys*, pp. 255–265 (2003)

16. Bi, Z., Da Xu, L., Wang, C.: Internet of things for enterprise systems of modern manufacturing. *IEEE Trans. Industr. Inf.* **10**(2), 1537–1546 (2014)
17. Coetzee, L., Eksteen, J.: The Internet of Things – Promise for the Future ? An Introduction. In: *IST-Africa 2011 Conference Proceedings*, pp. 1–9 (2011)
18. Borgia, E.: The internet of things vision: key features, applications and open issues. *Comput. Commun.* **54**, 1–31 (2014)
19. Miorandi, D., Sicari, S., De Pellegrini, F., Chlamtac, I.: Internet of things: vision, applications and research challenges. *Ad Hoc Netw.* **10**(7), 1497–1516 (2012)
20. Hwang, Y.M., Kim, M.G., Rho, J.J.: Understanding internet of things (IoT) diffusion: focusing on value configuration of RFID and sensors in business cases (2008–2012). *Inf. Dev.* **32**(4), 969–985 (2016)
21. Hevner, A.R., Gregor, S.: Positioning and presenting design science research for maximum impact. *MIS Q.* **37**(2), 337–355 (2013)
22. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a framework for evaluation in design science research. *Eur. J. Inf. Syst.* **25**, 77 (2016). <https://doi.org/10.1057/ejis.2014.36>
23. Fette, I.: The WebSocket protocol. *Internet Eng. Task Force, Request for Comments* **53**(9), 1–79 (2011). <https://doi.org/10.1017/CBO9781107415324.004>
24. Lubbers, P., Greco, F.: HTML5 web sockets: A quantum leap in scalability for the web. *SOA World Magazine*, Article (2010)
25. Da Xu, L., He, W., Li, S.: Internet of things in industries: a survey. *IEEE Trans. Industr. Inf.* **10**(4), 2233–2243 (2014)
26. Datta, P., Bhisham, S.A.: Survey on IoT architectures, protocols, security and smart city based applications. In: *8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–5. IEEE (2017)
27. Yaqoob, I., et al.: Internet of things architecture: recent advances, taxonomy, requirements and open challenges. *IEEE Wirel. Commun.* **24**(3), 10–16 (2017). <https://doi.org/10.1109/MWC.2017.1600421>
28. Lloret, J., Tomas, J., Canovas, A., Parra, L.: An integrated IoT architecture for smart metering. *IEEE Commun. Mag.* **54**(12), 50–57 (2016)
29. Kohavi, R., Quinlan, J.R.: Data mining tasks and methods: Classification: decision-tree discovery. *Handbook of data mining and knowledge discovery*. 267–276. Oxford University Press, Inc. New York, NY. USA (2002)



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





IoT Enabled Process Innovation: Exploring Sensor-Based Digital Service Design Through an Information Requirements Framework

Niclas Carlén, August Forsman , Jesper Svensson  ,
and Johan Sandberg 

Department of Informatics, Umeå University, 901 87 Umeå, Sweden
{niclas.carlen, august.forsman, jesper.svensson,
johan.sandberg}@umu.se

Abstract. Through digitisation of physical artefacts and environments, the Internet of Things carries vast potential for process innovation. However, navigation of the quickly evolving technological landscape and identification of emerging opportunities for value creation remains challenging. To this end, we combine existing frameworks on information requirements, IT capability, and business value of IT. We evaluate the usability of these frameworks for IoT enabled innovation in our analysis of two sensor-based process innovation projects. We investigate the fit between process characteristics and technological functionality, and the implications of this alignment. Our analysis demonstrates that the framework provides a practically useful and theoretically coherent conceptual device for analyzing process characteristics and digital options to innovate processes. Furthermore, we find that IoT sensors are well suited to address connectivity and uncertainty requirements. However, in order to leverage them to address high equivocality requirements designers need deep contextual understanding to align IoT capability with information requirements.

Keywords: Process innovation · Internet of Things · Information requirements

1 Introduction

The ongoing pervasive digitisation of physical artefacts and environments, often collectively referred to as the Internet of Things (IoT), signifies a new paradigm in data processing and communication. It is not a new technology per se, Wortmann and Flüchter (2015) describes IoT as “a global infrastructure for the Information Society, enabling advanced services by interconnecting (physical and virtual) things based on, existing and evolving, interoperable information and communication technologies”. IoT has evolved into an ecosystem of possible objects or things to connect to the internet where basically anything that has an on and off switch can be connected e.g. sensors. Combined with an expected battery life of several decades for many sensors allows for data collection around the clock. Sensors like this are making headway around the world with the emergence of smart cities, i.e. cities that are connected and efficient collecting data to manage its resources and serve its citizens in the best way

possible (Sundmaecker et al. 2010). The global market for IoT is growing tremendously, resulting in significant market opportunities and an expected turnover at around 8.9 trillion dollars combined with an estimated 26–100 billion connected devices by the year 2020 (Statista 2017). Currently, there are significant uncertainties when companies engage in IoT; lack of standards, low understanding of technology and security issues are all sizeable problems hindering development (Forbes Insight 2017). Most companies feel uncertain which IoT-solution will collect the data they need, this results in a discrepancy in what companies need and what they get with regards to data and information (Forbes Insight 2017). In this paper, we argue that this uncertainty can be mitigated by synthesising two theoretical frameworks combining both analysis of information requirements and corresponding digital options with effects on processes by different IoT implementations.

We have been involved extensively in two IoT-projects and have seen issues regarding information requirements and process analysis arise throughout both. Both cases illustrate the need to understand the working environment and the technology at hand. While most studies thus far have focused either on the technical or business side of IoT (Forbes Insight 2017), we argue that there is a need to understand both aspects to generate successful outcomes. Since in most cases there is going to be humans interacting with the technology in their daily work life, we also argue that it is of great importance to understand the duality of technology and social activity. We have, therefore, focused on a micro-level perspective by investigating how IoT-systems affect work practices. This paper seeks to examine the effects of implementing sensor-based systems in organisations, i.e. how the implementation effects the process and value chain. This study will be limited to two separate case studies where sensor-based systems have been both designed and implemented in the organisations. Specifically, we explore the following research question:

How can IoT sensors be used for digital service design to innovate processes?

2 Internet of Things in Business Processes

With the emergence of ICT in the 20th century a lot of IT projects have had lacklustre results, one reason is the assumption that technological functionality solves organisational issues (Alter 2006). The argument against this is that an IT system should be viewed as a resource for the solution and not the solution itself (Alter 2006; Meyer et al. 2013). IoT is also argued to have a decentralising effect on the value chain in business as parts of it become connected. As the value chain becomes decentralised the decision-making rights are moved to the individual components in the chain, and the different parts become more independent, arguably creating a more efficient and streamlined process chain throughout (Haller et al. 2008). We would add to the current research by enacting these principles in real-world implementations. Doing so with the intention of showing real-world scenarios with the technology in place, we argue that the current body of research will benefit by the holistic perspective applied in this study.

2.1 A Framework for Analysing Process Effects with IoT and IoT Capability

In this paper, we synthesise two different theoretical frameworks. The first is a framework developed to identify what digital options are available to organisations and what information requirements different processes have (Sandberg et al. 2014). The second framework is intended to measure the effects of the chosen solution on the process (Mooney 1995). By synthesising we refer to using the frameworks in tandem, iteratively, throughout the process analysis, implementation and evaluation of the systems. This in turn contributes to the body of research with a practical framework encompassing the entire process of implementing IoT in an organisation with intent to innovate business processes.

Information Requirements

Information requirements are used to identify what digital options there are for a specific business process. Digital options should here be understood as opportunities for leveraging IT in process innovation. The conceptualisation is grounded in earlier research, but Sandberg et al. (2014) further develop the concept of a specific tasks requirements about uncertainty and equivocality by adding connectivity to modernise the theory. We extend existing applications of the framework by using the information requirements in an IoT context. Establishing information requirements for each case is the first step of the innovative work process to map if and how IoT can support or transform a business process. This analysis, enables generation of the digital options available to organisations. To identify information requirements there are three aspects to consider and analyse: connectivity, uncertainty and equivocality.

Connectivity relates to the informational dependencies between processes and systems within an organisation, i.e. the need for information sharing across boundaries in an organisation (Malhotra et al. 2005). If the connectivity need is regarded as high, the focus for managers should be to counteract technical or social barriers by increasing information *reach*, characterised as “the number information sources that can be accessed during task execution” (Sandberg et al. 2014, p. 428). If information can be accessed across organisational and geographical boundaries, process information requirements have low connectivity requirements. When the connectivity need is low, it can be relevant to increase *richness* which refers to “the number of data points available regarding a given object during task execution” (Sandberg et al. 2014, p. 428).

Uncertainty refers to the availability and accuracy of information needed for actors to execute their task within an organisation. Uncertainty requirements can be addressed by continually balancing information production and information consumption (Ramaprasad and Rai 1996). Information *production* occurs when actors generate new information based on stimuli in the process and its environment. Information *consumption* turns the existing and available information into business process actions. If the information requirements are high in uncertainty, i.e. the current information is inaccurate, unreliable or insufficient, organisations should aim to increase production of information as they should not want to consume unreliable information. When uncertainty requirements are low, organisations can instead focus on consumption of

information. If the information is reliable, available options for consumption emerge which can drive an organisation to make better data-driven decisions.

Equivocality refer to confusion and lack of mutual understanding when executing a task, or the level of complexity and ambiguity in the tasks information processing (Sandberg et al. 2014) (Mathiassen and Sørensen 2008). When a task is to be executed, there may be a need for mutual understanding between actors or processes, as they may have to rely on contextual knowledge. Conversely, there can be situations where actors or processes can rely on their codified knowledge and routines. If information requirements are high in equivocality, then the situation or task context is unknown for the involved parties. A *relationship* enables success in these situations; there needs to be a high level of understanding and trust within the system and in the systems supporting situational knowledge and contexts. If the requirements are low, however, then the characteristics are based on *encounters* where standardised protocols and workflows are utilised.

Table 1. Information requirements and digital options characteristic (Adapted from Sandberg et al. 2014)

Information requirement	Corresponding digital option characteristic		Example of IoT capability investment
Connectivity	High	Reach: the number of information sources that can be accessed through IoT during task execution	Open data sharing of sensor data between organisational departments, generating easy access to new information
	Low	Richness: the number of data points available through IoT about a given object during task execution	Flow sensors generating exact measurements of waste flow in a waste management facility
Uncertainty	High	Production: the extent to which IoT supports the creation of information from stimuli	Multiple sensor measuring soil and crop health in modern agriculture
	Low	Consumption: the extent to which IoT support translation of information into action taken	Heatmaps showing movement patterns of visitors in a public building
Equivocality	High	Relationship: extent to which IoT supports contextual consideration and development of trust by adaptation and sharing of information across subsequent episodes	Correct analysis of contextual environment with the implementation of sensors that measure the exact values needed for the task
	Low	Encounter: IoT based on a standardised approach without variation across customers; limited regarding time and flexibility but efficient due to uniformity	Photoelectric sensors measuring visitors entering and leaving a facility from a permanently fixed position

IoT Capabilities in Organisations

Sandberg et al. (2014) refer to the organisation's IT capabilities as a firm's previous investments in IT resources, such as technology or IT competence. We use the concept and expand upon it to fit the field of IoT and thereby refer to them as IoT capabilities. Table 1 shows both information requirements, their corresponding digital options and examples of IoT capabilities that reflect the requirement characteristic.

2.2 Measuring Effects of IoT on Processes

We draw on Mooney's (1995) framework for assessing the business value of IT. Business performance is best measured from the performance of its processes (Ray et al. 2004; Mooney 1995). Sensors collect a limited range of data in an environment which in turn supports business processes. Therefore, we conclude that examining sensor implementation impact is most efficiently done in the context of the processes which they intend to support. In this study, we analyse the automational, transformational and informational effects of the IT-systems designed for each case, which are classifications derived from Mooney's process-oriented framework and have been used in similar studies (Stenmark and Jadaan 2010; Visich et al. 2009).

The automational dimension relates to how sensors data collection can substitute manual labour. Different kind of sensors can continuously collect data which are to support or initiate processes. Automational effects on business value are gained through aspects such as improved customer service, increased productivity and a more efficient labour distribution (Stenmark and Jadaan 2010; Visich et al. 2009).

Informational effects are those that are caused by IT-enabled collection, storage, processing and spread of information acquired from the sensors. Case studies of RFID implementations have shown that business value can arise from improved resource management and reduced manual labour (Stenmark and Jadaan 2010; Visich et al. 2009).

The transformational dimension affects and supports process innovation and transformation. Sensor data may support and improve existing processes but may also be utilised for business innovation. Data acquired to support a specific process chain can be used in combination with other aspects of the organisations' knowledge base to innovate the business (Stenmark and Jadaan 2010).

3 Methodology

For this paper we conducted a multiple case study consisting of two cases where sensor-based IoT-systems, through collection and visualisation of data, were designed and implemented to support different process chains. Case studies are a preferred strategy when research questions related to "how" and "why" are posed (Yin 2003), and multiple case-studies when the logic of the study is to "produce contrasting results but for predictable reasons" (Yin 2003, p. 47). We argue that this makes a case-study approach viable with regards to the framing and research question stated earlier. By evaluating this innovation process through multiple cases, we intend to generate general findings and propose practices which could be built upon in further research.

We chose the cases based on an analysis of two different research sites. The classification for inclusion was that the case should present one or several concrete problems in a process chain where sensor-based technology could be a solution. These problems could either be a lack of ability in performing activities which could be enabled by the technology or addressing problems currently present in an organisation.

A process is “a structured, measured set of activities designed to produce a specified output for a particular customer or market” (Davenport 1993, p. 5) and can be classified into two different categories; operational processes and management processes (Mooney 1995). Operational processes are the set of activities an organisation performs to produce something that generates value and is referred to as an organisation’s primary activities. Management processes are related to streamlining and improving the efficiency of an organisation’s primary set of activities such as coordinating and handling different information. Process innovation in this context refers to the practice of analysing an organisation’s processes and redesigning them using innovative technology to improve performance and support the processes (Davenport 1993). In this case, that innovative technology is LoRa-sensors, which enables remote monitoring and control of different aspects of a process, and the IT-artefacts of the software designed to visualise or manipulate the data generated by the sensors. Each case process chain was broken down into sub-processes depending on what type of activities and the complexity of the tasks performed. We then mounted sensors at each research site to collect data and designed IT-artefacts with the purpose of solving specific problems related to the sub-process. The effectiveness of these systems was analysed in the context of what type of value and effects the data generated when innovating operational and management processes.

3.1 Research Sites and Sensor Technology

The sensors used in the project are based on LoRa technology, and a specific type of gateway delivers connectivity to the sensors. Because of lacking infrastructure in the northern municipality, we were tasked to mount two base-stations; this implementation took place in September 2017. One of these base-stations was installed on the highest point in the town, which is a water tower, the other one on the roof of the local secondary school. We also installed a base-station in the city where the cleaning company project took place to further the existing coverage. The base-stations have guaranteed coverage of a 3 km radius, however, depending on disruption and quality of air it may be greater than that (LoRa-alliance 2015).

LoRa stands for Long Range and is the physical layer utilised to create long-range communication links. LoRa is based on chirp spread spectrum modulation, using its entire spectrum of bandwidth and is therefore very resistant to channel noise. LoRaWAN is based on Low Power, Wide-Area Networks i.e. LPWAN, and defines the communication protocol and system architecture for the network. LoRaWAN has a great influence on battery times for nodes, network capacity and security. LoRaWAN is explicitly designed for sensors and applications that need to send small amounts of data over long distances at different time intervals, making it ideal for IoT sensors applications. As IoT is still a new phenomenon and standards are currently lacking, other technologies are competing to become the business standard. In likeness with LoRa,

several LPWAN networks are emerging as competitors; examples are Sigfox and Narrowband IoT that operate similarly (LoRa-alliance 2015).

In this study, we have used two different types of LoRa-sensors that measure different values while having some similar readings; temperature inside the casing, humidity at the sensor and battery-level (Table 2).

Table 2. LoRa Sensors utilised in projects with corresponding properties

Sensor name	Properties
ERS	Passive infrared (PIR) sensor registering movement in its field of view
ELT-1	Analogue input sensor capable of coupling with external analogue measurement tools, e.g. thermometer, voltmeter, ultrasonic level indicator

3.2 Data Collection and Analysis

The data collection process consisted of an analytical phase where we in conjunction with each organisation collected data regarding the problem background of their process chains. An implementation phase studying the physical environment of the research site, collected data relevant to the practical implementation of the system, and an evaluation phase where the data related to the results of the system was collected. All interviews conducted are semi-structured and were conducted throughout all three phases to capture viewpoints from the participants on both current problematic aspects, initial impressions of the implemented system, and impressions regarding its effects. Observations refer to activities where we have studied staff members performing the process chains, as well as documenting areas of interest. This was performed mainly during the analytical phase in order to discover details relevant to mapping the process chains and capture insights possibly missed using the other methods. Workshops are meetings where we collected data related to our subsequent design choices and was mainly conducted in the analytical phase with participants in each innovation project. Informal encounters are the interactions with the organisation where we have performed different tasks or exchanged minor pieces of information related to the projects, a method which was applied throughout all three phases.

A total of 9 interviews, 11 observations, 7 workshops and 46 informal encounters were conducted, spanning all three cases in the study during a time frame spanning from 20th September 2017 to the 29th of March 2018 (Table 3).

Table 3. Data collection overview

Case	Interviews	Observations	Workshops	Informal encounters
Swimming pool	4	5	2	7
Cleaning company	5	6	5	39

The data analysis was performed in iterations together with the organisation throughout the time-frame of the data collection, where the input in the analytical phase formed the basis for our description of the information requirements and IT-capabilities in each case. The results of the evaluation period (testing the systems in practice) formed the input for the process effects each system had on the corresponding process chain it supported.

4 Results

This section consists of the results of our research, each case will be presented with a description of the research site, problem background and information requirements of the process chains, with a subsequent description of practical implementation and effects from each case.

4.1 Swimming Pool

Research Site

The research site of the public swimming pool-case is in a municipality in the north of Sweden and is run by a small organisation of six people who maintain the pool, a gym and a gymnasium in the same building. The public pool is a facility open during weekdays and Saturdays on regular weeks.

Process Chains and Problem Backgrounds

The organisation lacks data on the number of visitors and which hours and days during the week generate most activity. The main areas of activity which the organisation found interesting were the entrance to estimate the overall number of customers, the cafeteria to investigate the air quality during peak hours, and the locker rooms to investigate differences in attendance between the genders. This data is interesting for the organisation when optimising staffing and air quality, and to create an overview of when and how much the facilities are used during the week. To achieve these informational effects, the sensor implementation sought to address the high uncertainty and low connectivity information requirements through continuous collection of visitor data. Further, the information regarding the number of visitors is non-equivocal as the collected data is readily interpreted in the context.

The second process chain in the case is optimising heating of the swimming pool. According to the person responsible for this routine, the pools are heated to 32 °C every week during Tuesday nights, and then the temperature falls successively to around 27 °C during the weekly cycle. This practice leads to uncertainty amongst the customers on the current temperature and generates phone calls to the organisation increasing the workload. Further, the facilities have shown signs of increased wear in forms of mould and moisture damage due to the increased evaporation generated by higher temperatures. The organisation seeks informational effects on the managerial level through an increased amount of temperature data points and implementing sensors to address the connectivity and uncertainty requirements. The equivocal requirements were low as temperature data is readily understood in the context.

The third process chosen for this case is the documentation of pool water quality, which is a process chain performed by the staff daily to discover anomalies and potential health risks related to the pool water. Water samples are collected, analysed and documented as the first task of every day: water temperature, pH-value and chlorine-levels. To collect this data, the staff places a thermometer in the pool water where it is submerged for 15 min. During this time, they gather two water samples which are analysed using a pool water quality kit establishing its pH-value and amount of chlorine. This data is then documented manually in a binder and stored in the staff office of the facilities. Due to the repetitive manner of the data collection and documentation, the organisation seeks to explore to what extent it could be automated using IoT sensors. To achieve these effects, automatic production of information regarding temperature, pH-value and chlorine levels is required, which characterises an information requirement high in uncertainty. We found the connectivity and equivocality requirements to be low as the information is to be utilised within task entity boundaries and well understood in the context.

Implementation

To support the first process chain, we mounted four sensors at areas for which the organisation had expressed interest. The main units of observation these sensors were to measure were motion activity and temperature in each respective area. The sensors were placed at the entrance of each respective area at the height of around 150 cm's ensuring measurements of every individual passing.

This motion activity and temperature-data is uploaded every 30 min to a database, imported into tables and transformed into graphs, both real-time and historical. The information was made accessible to the organisation through a web application where it could be studied and form part of the basis for process innovation. The implemented system has the characteristics referring to the production of information to address uncertainty requirements. As this process aims to collect information about motion and temperature in the facility we argued that the four sensors would address the uncertainty requirements.

To support the second chain, we mounted one temperature sensor in the bottom of the swimming pool, hidden behind a ladder. This sensor uploaded water temperature every 30 min to a database and was imported into tables and graphs made available to the organisation. The historical data generated by the sensor can be utilised to measure how much time it takes to heat the pool to the preferred temperature and get a more detailed overview of its heating cycle. This data could serve as a basis for innovating the heating process chain and minimise the problems of their current practices. The third process chain utilised the same temperature sensor as the second as the only relevant unit in the process chain collected and documented is the temperature data. This data was then uploaded every 30 min and presented in the form of tables and graphs. For this process, we addressed the information requirements with the intent to increase richness by collecting data with the sensor as well as produce more information to lower uncertainty.

Effects

The first process chain showed primarily informational effects. The data collected from the motion sensors generated an estimate of which areas has the most activity, and

during which hours can be established by studying the graphs in the web application. The organisation had ideas of using the data to optimise the air conditioning. However, the functional capabilities of adjusting the air conditioner cycles to reflect usage or be automated by the data seemed to be limited which was uncovered later in the project when this issue was discussed with a janitor responsible for the air conditioning. A side-effect of studying the comprehensive dataset was the discovery that the temperature in the cafeteria rises around 2–3 °C during the nights when the facility is closed. This occurrence was unknown to the organisation when they were informed of it, and according to one of the staff members, may be related to the underfloor heating being active during the night time when the ventilation is inactive.

The effects on the second process chain had informational effects. The historical data generates a clear and consistent timeframe over how long it takes for the swimming pool to reach its intended temperature and shows some anomalies. An example is the re-warming of the pool, which usually happens around 4 h after it has reached its maximum temperature. Why this happens is unknown to us right now but will be of interest in further evaluation of the system. The third process chain has potential to be completely automated, generating both automational and transformational effects, but since the sensors automate only 1/3 of the data collection, the staff must still perform a majority of the process chain in the same manner as before. A future update of the system will be to implement sensors collecting data of the chlorine level and pH-value. With a complete system in place, the whole process will be performed continuously and automatically document the data in the same way as current practice. The permanent character enables transformational possibilities in the sense that with a system documenting the water quality continuously, anomalies in the water can be discovered faster.

4.2 Cleaning Company

Research Site

The research site for the case of the cleaning company is located in a university building. The organisation is responsible for cleaning all facilities and have a staff of six managing and executing this task at the research site.

Process Chains and Problem Backgrounds

The staff have expressed a problem of prioritising which order that classrooms are cleaned after the weekend in an effort to work more condition-based. The current situation is such that the activity in the classrooms during the weekends is unknown to the staff when they begin cleaning on Mondays, and they clean each room in a set routine. The consequence of cleaning in a routine-based manner could result in rooms with less cleaning needs getting cleaned, and rooms with higher cleaning need left unattended. The cleaning process of each room constitutes four sub-processes: cleaning the floor, cleaning the tables, wiping the whiteboard and emptying the waste bin, and the cleaning need of each sub-process in combination is what constitute the cleaning need of the classroom. The organisation seeks informational effects through collection of information regarding classroom cleaning need. This information requirement has the characteristics of high equivocality, as assessing cleaning need is non-algorithmic

and is based on situated knowledge about the specific context. The process chain has high connectivity and both high and low uncertainty requirements as cleaning need information must be collected and accessed remotely. Since the organisation currently lacks data on the activity in the classrooms during the weekends, it cannot innovate its processes in such a way that it aligns with the ambition of working more condition-based.

The second process chain presented as problematic by the organisation is assessing if a room is vacant. In the current situation, the staff do not clean rooms which are occupied and wait until the rooms are vacant to clean them. This problem means in practice that they sometimes spend time visiting rooms only to discover that they cannot be cleaned, and delay that process until later, having wasted time moving to the classroom. The organisation seeks informational effects on the operational level through remote access to information with regards to classroom vacancy. To achieve these effects, vacancy information need to be produced and remotely accessible to cleaning staff, which characterises the information requirements as high in connectivity and uncertainty. Lastly, information regarding classroom vacancy is non-equivocal as the room is either vacant or not.

The third process chain is the comparison between the presumed usage based on the booking schedule and actual usage of the specific classrooms. According to the organisation, it is not uncommon for a room to be booked during the week, but its actual usage is unclear. The staff can plan the cleaning of classrooms only to discover that they have not been used and, therefore, not in need of cleaning. This could also be used as a basis when negotiating terms with its currently largest customer which is the university itself. Part of how many hours the company can bill the university is based on the number of hours booked in the electronic booking schedule. The organisation seeks informational effects on the managerial level through remote information collection of classroom usage. High connectivity, high uncertainty, and low equivocality characterises its information requirement. The multiple data collection points increase reach and production of this non-equivocal information.

Implementation

To support the process chains described in the case we mounted sensors in classrooms collecting motion data. The sensors were placed at the entrance around 170 cms from the floor, registering every motion near the entrance door. This data was uploaded every 10 min to a database. We then designed a web-application containing various artefacts which utilise this data to address the information requirements in the process chain. Due to the high connectivity requirements the purpose was to increase reach through multiple data collection points. For the uncertainty requirements, there was a need for both production and consumption to address the requirements relating to uncertainty. The sensors installed addressed the information production aspect, and the web application was developed to increase consumption of information. As the equivocality requirements were high, the need for a relationship characteristic was of high priority. Due to the organisation wanting to measure cleaning need in the rooms, which is a highly equivocal measurement, the sensors and web application needed to be utilised in conjunction with the cleaning staffs' knowledge and routines.

The artefacts contained in the application display three sets of data; accumulated motion in each room, a two colour-button signalling if motion has been detected the last 10 min and historical data available for export in the form of graphs. The application was made accessible for the cleaning staff in their day-to-day work by a tablet placed on their cleaning cart.

The system was tested for three weeks, during which the staff had access and utilised it when performing their tasks. Furthermore, they graded the experienced cleaning need which was defined on a three-grade scale where one was clean, two was normal and three related to a high cleaning need. Interviews with the staff were conducted before, during and after the test period. During this period, we also tested the hypothesis that an increased motion value from a classroom during the weekend represents a higher cleaning need. This hypothesis was tested by photographing every aspect related to the sub-processes of the cleaning process chain after the weekends, comparing the empirical findings with motion data captured by the sensors.

Effects

The implemented system had various effects on the organisation depending on which process chain it supported, but how well it improves the general organisational performance remains inconclusive and needs to be evaluated further. Although the system generated the motion data we presumed when designing and implementing the system, the usability of this data in the context of the first process chain, determining the cleaning need of a specific classroom, is still unclear. The system was designed to have mainly informational effects on the first process chain, by presenting information to the staff which could be used to determine which classrooms that had a more significant cleaning need. The empirical findings, however, related to the hypothesis that higher motion value represents a higher cleaning need are vague. We believe that further evaluation of the system is necessary to establish its effects.

The effects on the second process chain, which was to inform the staff if a classroom is vacant and possible to clean, has mainly been of informational and automational character, with the intended ability successfully generated. According to the staff, the system correctly identifies if a classroom is vacant or if there are students present, which informs them in a way that improves their performance. Since they do not have to spend time collecting this information manually, it has automated the process chain. The long-term effects of having this ability, its possible flaws (no motion input if students are very still) and how great of a value it brings to the organisation, due to it solving a relatively minor problem, will have to be further evaluated.

The system has generated the desired ability to analyse and compare between the booked hours in the electronic booking schedule and the actual amount of activity in the classrooms, which would classify it as having a transformational effect on the organisation's ability to innovate. This ability may contribute to having informational effects which improve performance, depending on how generated data is utilised. The organisation has expressed an ambition to integrate the graphs, and sensor data with the current electronic booking schedule to easier compare the data, but this feature has yet to be implemented.

5 Discussion

To provide actionable guidance for the use of IoT sensors to innovate processes through digital service design, we have illustrated the applicability of a synthesised framework facilitating opportunity recognition, design and analysis of effects. The analysis provides insight both to the general applicability of the framework across the innovation process, and the bearing of specific components of the framework for IoT sensors.

Although developed for different tasks, our application of the framework suggests that it is beneficial to apply the whole chain of analysis in the different subparts of the innovation process. While the business value of IT provides support for retrospective analysis of effects (Stenmark and Jadaan 2010; Visich et al. 2009), the desired outcomes in terms of informational, automational and transformational effects should guide the design of the digital service system. Thus, in accordance with Alters (2006) arguments regarding a holistic view in systems design, such ambitions need to be considered in the initial analysis phase. By establishing information requirements for a process chain or specific sub-processes during this phase, potential complexities related to generating the desired effects can be discovered, e.g. processes with a high level of equivocality. With a desirable effect-outcome and the information requirements necessary for generating this outcome established, we argue that this provides a more well-grounded basis for process innovation with sensors.

For sensor-based process innovation, the process information requirements in part determine the degree of automational effect that can be achieved. A process with the purpose of simply collecting or communicating one type of data may be automated in its entirety through implementing a sensor-based system. Automational effects of this kind are shown in the swimming pool case where the process chain of documenting pool water quality has the potential of being automated completely by the utilisation of sensor-based systems. This high degree of automational effects arises from alignment between information needed to complete the task and sensor capacity to produce data output. This outcome differs from a process with high equivocal information requirements, such as the process of evaluating the cleaning need where there are four sub-processes to complete the process chain. Each sub-process requires information with regards to its specific cleaning need, and the sum of the informational output from these sub-processes are then what constitutes the cleaning need of the classroom. The sensor used to support this process chain provides information of movement around the entrance to the classroom. This information of movement does not map precisely to any of the sub-process outputs of assessing a cleaning need. For example, to assess if a whiteboard needs cleaning there is a requirement of visual examination of the whiteboard and from this draw a conclusion regarding its need for cleaning. Information of movement is only a proxy variable that does not directly respond to use of the whiteboard. Similarly, to assess if the floor needs mopping, information is required on the amount of dust and dirt that is currently present on its surface. Again, the type of information provided by the sensor used does not immediately support an assessment of mopping need.

The sensor-produced data may be used to make assumptions of cleaning need based on the information of activity around the sensor, without showing a one to one relationship between the amount of movement and the equivocal cleaning need. For

instance, while movement information does not show how dusty or dirty the floor is, it shows the cause of this effect: people have walked on the floor. Any value of activity data means that the floor has been walked upon. If the floor has been walked upon, it is reasonable to assume that some amount of dirt and dust have been transferred from the shoes to the surface of the floor. Thus, a value of activity data increases the need of mopping the floor.

This reasoning may be used on the other sub-processes of the cleaning case as well, though with a weaker conclusion. We have argued that movement information show a relationship with the degree of dirt on the floors. However, movement information is not directly correlated with the use of whiteboards. Recorded activity data means that there has been a person around the entrance of the classroom. It does, however, not capture the type of activity the person has engaged in, e.g. if they used the whiteboards. Similarly, the data does not show if the person(s) moving close to the classroom entrance also throw waste into the waste bin. The IoT sensor capabilities do not address these equivocal information requirements directly. Thus, conclusions regarding the cleaning need of these processes cannot be drawn solely from the information provided by the sensor. Movement information may, however, be used in conjunction with visual evaluation of the state of classrooms over time to show statistical probabilities of whiteboard cleaning need, waste bin level and table dirt. Cleaning staff collected this data during the testing phase of the cleaning case by grading the classrooms total cleaning need in conjunction with them cleaning it. This method of evaluation could be improved by splitting it into an evaluation of each sub-process, thus increasing its accuracy. There are degrees to this relationship between process information requirements and sensor data output. This relationship spans from misaligned, exemplified by assessing cleaning need of whiteboards with movement sensors, to aligned, as shown by addressing temperature requirements with a temperature sensor. Thus, we argue that IoT sensors could be implemented to processes with various degrees of success depending on the process equivocality information requirements (Table 4).

Table 4. General descriptions of IoT capabilities addressing information requirements.

High connectivity <i>Reach</i>	IoT-sensors by nature increases reach due to many individual data collection points that are accessible over organisational borders
High uncertainty <i>Production</i>	IoT-sensors can continuously produce data outside human intervention, day as night
High equivocality <i>Relation</i>	IoT-sensors can supply relations, however, important to match sensor capabilities with the process information requirements
Low connectivity <i>Richness</i>	IoT-sensors add richness and granularity due to continuous, focused data collection
Low uncertainty <i>Consumption</i>	IoT-sensors do not directly address consumption, but through data-analysis and visualisation of sensor-data consumption can be addressed
Low equivocality <i>Encounter</i>	IoT-sensors can fortify encounters through stabile data collection that supports codified knowledge and routine-based tasks

6 Conclusion

In this paper, we provide a synthesised framework for IoT sensor-based process innovation. The framework draws on extant theory on (1) the role of information requirements analysis in identification of IT-based process innovation opportunities and (2) effects on business value. The synthetisation of these theoretical devices enables a holistic analysis, from opportunity recognition to evaluation of achieved effects. We have explored the practical usability of the framework through an analysis of two different implementations of IoT systems and associated organisational effects. We found areas where the sensors have shown potential for process innovation and demonstrated the applicability of an information requirements perspective in an IoT context. Through this we have been able to identify areas where complexity becomes an issue for implementation of IoT systems. In particular, the results point to IoT sensors general capacity for responding to high connectivity and uncertainty requirements, and the need for aligning functionality with organisational needs to respond to high equivocality requirements. Thus, the functionality provided by IoT sensors does not reduce the importance of organisational ability for process analysis and identification of values to measure. We also identify the effects and relate them to levels of alignment between information requirements in the process and sensor capacity. Lastly, we have argued for a need to understand IoT through a micro-level perspective in organisational processes and proposed a set of practices for overcoming challenges encountered in the cases. As the field continues to grow we believe it is essential for organisations to further understand how to utilise IoT-technology when innovating their work processes.

References

- Alter, S.: The Work System Method, 1st edn. Work System Press, Larkspur (2006)
- Davenport, T.H.: Process Innovation - Reengineering Work Through Information Technology, 1st edn. Harvard Business School Press, Boston (1993)
- Insight, F.: Internet of Things - From Theory to Reality. Forbes Insight, Jersey City (2017)
- Haller, S., Karnouskos, S., Schroth, C.: The Internet of Things in an Enterprise Context. Springer, Vienna (2008). https://doi.org/10.1007/978-3-642-00985-3_2
- LoRa-alliance: LoRaWAN - What is it? A technical overview of LoRa and LoRaWAN, San Ramon, CA: LoRa® Alliance Technical Marketing Workgroup (2015)
- Malhotra, A., Gosain, S., Sawy, O.: Absorptive capacity configurations in supply chains: gearing for partner-enabled market knowledge creation. *MIS Q.* **29**, 145–187 (2005). <https://doi.org/10.2307/25148671>
- Mathiassen, L., Sørensen, C.: Towards a theory of organizational information services. *J. Inf. Technol.* **23**, 313–329 (2008). <https://doi.org/10.1057/jit.2008.10>
- Meyer, S., Ruppen, A., Magerkurth, C.: Internet of Things-Aware Process Modeling: Integrating IoT Devices as Business Process Resources. Springer, Valencia (2013). https://doi.org/10.1007/978-3-642-38709-8_6
- Mooney, J.G.: A Process Oriented Framework for Assessing the Business Value of Information Technology. Center for Research in Information Technology and Organizations, University of California, Irvine (1995)

- Ramaprasad, A., Rai, A.: Envisioning management of information. *Omega*. **24**, 179–193 (1996). [https://doi.org/10.1016/0305-0483\(95\)00061-5](https://doi.org/10.1016/0305-0483(95)00061-5)
- Ray, G., Barney, J., Muhanna, W.: Capabilities, business processes, and competitive advantage: choosing the dependent variable in empirical tests of the resource-based view. *Strateg. Manag. J.* **25**, 23–37 (2004). <https://doi.org/10.1002/smj.366>
- Sandberg, J., Mathiassen, L., Napier, N.: Digital options theory for IT capability investment. *J. Assoc. Inf. Syst.* **15**(7), 422–453 (2014)
- Statista: Statista - The statistics portal (2017) <https://www.statista.com/statistics/512673/worldwide-internet-of-things-market/>. Använd 13 March 2018
- Stenmark, D., Jadaan, T.: Enabling process innovation through sensor technology: a multiple case study of RFID deployment. In: *ECIS 2010 Proceedings*, Gothenburg (2010)
- Sundmaeker, H., Guillemin, P., Friess, P., Woellflé, S.: *Vision and Challenges for Realising the Internet of Things*. Publications Office of the European Union, Luxembourg (2010)
- Visich, J.K., Li, S., Khumawala, B.M., Reyes, P.M.: Empirical evidence of RFID impacts on supply chain performance. *Int. J. Oper. Prod. Manag.* **29**(12), 1290–1315 (2009)
- Wortmann, F., Flüchter, K.: Internet of things - technology and value added. *Bus. Inf. Syst. Eng.* **57**(3), 221–224 (2015)
- Yin, R.: *Case Study Research*, 3rd edn. SAGE Publications, Thousand Oaks (2003)



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





An Internet of Things Based Platform for Real-Time Management of Energy Consumption in Water Resource Recovery Facilities

Mário Nunes¹ , Rita Alves², Augusto Casaca¹ , Pedro Póvoa²,
and José Botelho²

¹ INOV/INESC-ID, R. Alves Redol, 1000-029 Lisbon, Portugal
mario.nunes@inov.pt, augusto.casaca@inesc-id.pt

² Águas do Tejo Atlântico, Fábrica de Água de Alcântara,
Av. de Ceuta, 1300-254 Lisbon, Portugal
{rita.alves,p.povoa,j.botelho}@adp.pt

Abstract. The article describes the design of an Internet of Things based platform having as main objective the real-time management of energy consumption in water resource recovery facilities and their integration in a future demand side management environment. The monitoring of several electrical parameters, including energy consumption, is done via a dedicated energy meter, whose design is detailed in the article. The high level data communication from the energy meters to a central platform of the wastewater utility is done via the MQTT protocol. Within the water resource recovery facility, the access network is based either on Wi-Fi or LoRa, which are two enabling technologies for the Internet of Things. The meters are deployed in pilot demonstrators located in two water resource recovery facilities in Lisbon, Portugal.

Keywords: Wastewater management · Internet of Things · Smart cities · Energy meters

1 Introduction

The Internet of Things (IoT) is the support for a large number of applications. Many of these applications are fundamental for implementing the concept of smart cities, namely in domains like e-health, smart buildings, transportation systems, energy grids and wastewater management. In this paper we focus into a smart wastewater management application, supported in an IoT infrastructure, having in view the real-time control of energy consumption in Water Resource Recovery Facilities (WRRF) of a Portuguese Wastewater Utility.

Energy is often the second-highest operating cost at WRRF after the labor cost; it can be above 50% of an utility's total operating costs [1]. The reason for this is that

most of the processes that occur in WRRF require energy for their operation and are intensive energy consumers. Also, when the drainage system is not gravitational, energy is required for operating pumping stations. In Portugal, the Águas de Portugal (AdP) Group, which is the utility in charge of water distribution and wastewater treatment plants in Portugal, represents about 1.4% of the total electrical consumption of the country and the energy cost represents 60% of the operating costs of AdP.

On the other hand, nowadays, new challenges appear due to the limited water and nutrient resources, to the existence of the circular economy framework and to climate changes concerns associated with the fossil fuel consumption; additionally there is an increasing cost of energy for the utility. In this context, a new paradigm for the use of the domestic wastewater was created. Domestic wastewater is being looked more as a resource than a waste and the wastewater treatment plants now are known as WRRF, where it is possible to recover nutrients and water to achieve a more sustainable use of the wastewater energy potential, and become a driver for the circular economy [2].

Águas do Tejo Atlântico (AdTA), which is one of the companies of the AdP group, is in charge of the wastewater treatment in the Lisbon region and in the west region of Portugal. AdTA is a consortium member of the running European Union H2020 research project “Intelligent Grid Technologies for Renewables Integration and Interactive Consumer Participation Enabling Interoperable Market Solutions and Interconnected Stakeholders”, which has the acronym INTEGRID. The main objectives of INTEGRID are to test the flexibility of electrical energy consumption for domestic and industrial consumers, to test energy storage systems and make forecasts of renewable energy production and consumption. The main role of AdTA in this project is to manage the flexibility of its internal processes in order to minimize the energy costs according to the market and to the requirements of the grid operators, leading to an optimization of the AdTA internal processes.

The AdTA focus on the flexibility of energy consumption and in the new challenges of turning a wastewater treatment plant into a WRRF is an important path towards its objective of achieving operational optimization and flexibility of processes. To optimize and make processes more flexible it is essential to know the performance of the processes. Thus, one of the variables to monitor is the energy consumption of the process or equipment. This paper is related with the specific work done for real-time monitoring of the electrical energy consumption in WRRF. It is being developed by AdTA, within the INTEGRID project, in collaboration with INOV, a Portuguese research institute.

The objective of this collaboration is to design and implement a low cost smart energy meter capable of being integrated in an Internet of Things (IoT) environment. The meter will periodically measure electrical energy consumption and several other electrical parameters, and will communicate those measurements to a central system and a database. The meter has a bi-directional communication with the central system, being allowed a remote control of the meter, namely for altering the configuration parameters. The meter is equipped with a state of the art communication technology compatible with the IoT communication protocols. Low cost for the meter is a must, as a large number of meters will be required for the complete universe of WRRF. Pilot demonstrators are being deployed in two WRRF in Lisbon, Portugal. The demonstrators have a total of 30 energy meters, and two different IoT communication

technologies were chosen to be tested in the two pilots. Each of the pilots also follows an IoT deployment framework with respect, not only to the communication protocols, but also to the used platforms.

The article is organized as follows. In Sect. 2 related work is reviewed, and in Sect. 3 the communication architecture is presented. Section 4 deals with the meter structure. The pilot deployment is treated in Sect. 5 and the last section concludes the paper.

2 Related Work

In the context of the ability to collect, transmit and process data, with a view to make the most of the collected information by transforming it into knowledge, there was a need at AdTA to implement systematic and integrated approaches, i.e., the implementation of decision support tools. Thus, the AQUASAFE platform was developed at AdTA [3], integrating data already existing in the company, in order to produce answers to the specific needs of the operation and management. The AQUASAFE platform is a structure that allows managing all the information flows necessary to obtain an adequate response in the context of the management problems (overflows, energy management, emergency response, etc.). The measured data, mainly from water flows, is imported in real-time and the models run periodically in the forecast mode in simulation scenarios chosen by the user. The AQUASAFE platform is in full use in AdTA nowadays.

With regard to energy use, energy systems are sensitive to energy consumption spikes and, therefore, measurements have to be taken, either to optimize energy generation and distribution or better to reduce or shift peak power demands. While there is plenty of experience in optimizing energy generation and distribution, it is the demand side management that is receiving increasing attention by research and industry [4]. Thus far, there still exists a gap between energy consumption and costs since there is no generalized cost model describing current energy tariff structures to evaluate operating costs at WRRF [5]. In most energy studies, the energy consumption is multiplied by an average energy price. However, operating costs significantly depend on the energy tariff structure applied. Different time-of-use and/or peak penalty charges may change the cost efficiency of a control solution completely [6].

For the first time, an application of a real energy pricing structure was applied to a calibrated model for evaluating operational strategies in two large WRRF, in the context of the Portuguese “SmartWater4Energy” project [7]. The importance and need of mathematical modelling for energy optimization of specific energy costs at real WRRF processes was assessed. Time periods with potential for further optimization were identified, supporting a smart grid basis in terms of water and energy markets that respond to the demands [6]. This work was developed in the AQUASAFE platform, where the different models and the data from different sensors (flow, energy consumption, dissolved oxygen concentration, NO₃ concentration, etc.) were included for calibrating and evaluating the models. However, in the AQUASAFE platform, the energy consumption measurements are done in an indirect way, through the SCADA system.

There is, however, a need to have accurate measurements on the energy consumption and other electrical parameters in real-time for all the WRRFs. Also, those measurements need to be communicated to a central platform with analytical capability, which allows extracting information from the data being measured. Based on the analytical studies performed at the platform it will be possible to devise a strategy for shifting peak loads and reducing energy consumption on the whole. These are the main reasons that originated the current developments described in this paper.

On what concerns the development of dedicated energy meters to measure several electrical parameters like current, voltage and power in real-time, and adapted for wireless sensor network communications, there is previous work already done for the smart grid environment [8, 9]. In the present case, the energy meters will be different from those ones as they must have the following distinctive characteristics: (i) to be adapted to the constraints of a WRRF deployment; (ii) to measure the parameters required in WRRF; (iii) to comply with the IoT communications paradigm and platform architecture, which is the state of the art for smart cities deployment; (iv) to be low cost.

3 Communication Architecture

The basic design idea for the communication architecture is to consider each meter, and the respective equipment to which it is connected, as a “thing” in an IoT context and collect the information from all the “things” in a central platform of AdTA, where the data can be stored in a database and data analytics be performed. Secure communication is a must and for that purpose we have available the AdTA private communication network that provides a security guarantee for the wide area communication.

The first decision to be taken was concerned with the IoT communication protocol(s) to use. From the universe of IoT communication protocols [10], we have considered that Wi-Fi and LoRa are two appropriate standard communication technologies for use in this solution and, therefore, in the pilot demonstrators. Wi-Fi is a well - known technology, having low cost communication modules for the meters, a low cost Access Point (AP), and high data rates (Mbps). As low cost is a key objective, the choice of Wi-Fi as one of the candidate technologies looked promising.

LoRa was the second candidate technology selected for the tests. LoRa is the physical layer containing the wireless modulation utilized to create a long range communication link. The complete stack of protocols used over LoRa is known as LoRaWAN. Compared to Wi-Fi, LoRa is a higher cost technology, namely for the LoRaWAN gateway, and has lower data rates (Kbps). However, it enables a longer communication distance than Wi-Fi, which is very useful for the communication between some remote equipment in the WRRF and the LoRaWAN gateway. The decision for the pilot demonstrations was to test both technologies, Wi-Fi in one WRRF (Chelas WRRF) and LoRa in the second one (Beirolas WRRF).

As it is required to transmit the data from the meters to the Control Centre, we had to establish a communication architecture that allows a seamless and secure transmission. The chosen architecture is shown in Fig. 1 for the Wi-Fi access case.

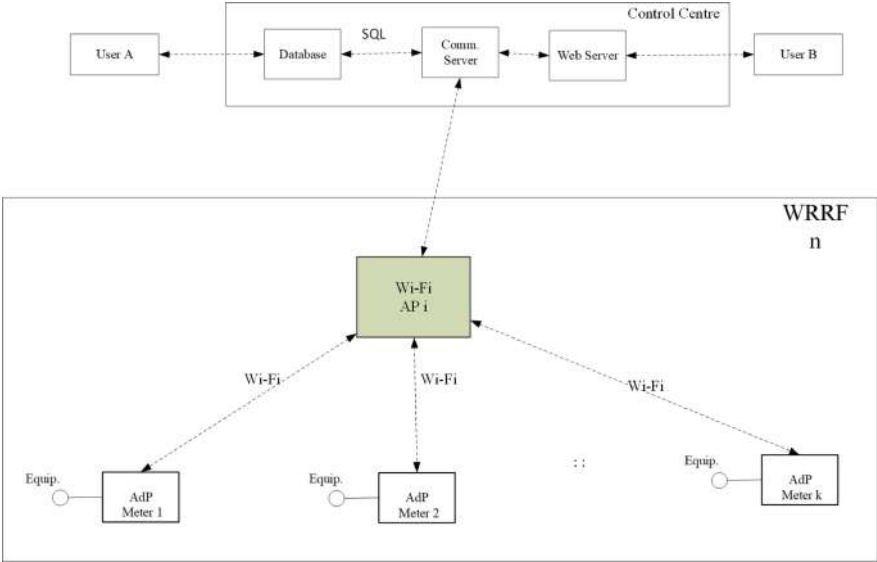


Fig. 1. System communication architecture based on Wi-Fi access

There might be several Wi-Fi APs installed in a WRRF. The WRRF energy meters, from now on called AdP meters, are deployed at the different WRRF equipments, e.g., recirculating pumps, equalization pumps, ventilators, etc. The meters communicate to the nearest Wi-Fi AP, sending a message containing the meter identification, followed by the electrical measurements. The data is forwarded from the Wi-Fi AP to the Communication Server (CS) located at the Control Centre via the AdTA private communication network. The CS will upload the data into the database, by means of the SQL protocol. The users can access the data either via a direct connection to the database or indirectly through a Web server.

Figure 2 shows a simplified protocol stack of the Wi-Fi based access network. A conventional TCP/IP stack of protocols is used over the Wi-Fi Medium Access Control (MAC) and Physical (PHY) layers. The Network server is implemented as one of the components of the CS in the Control Centre. The Wi-Fi AP converts Wi-Fi into Ethernet and communicates with the CS via the AdTA private wide area network.

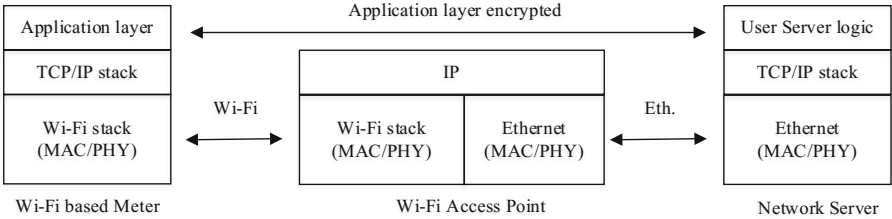


Fig. 2. Protocol stack of the Wi-Fi based access network

For LoRa, the system communication architecture is similar to the one shown in Fig. 1, having as main difference the use of a LoRaWAN Gateway instead of the Wi-Fi AP.

Figure 3 shows the protocol stack of LoRaWAN, which comprises 4 layers: RF layer, Physical layer, MAC layer and Application layer.

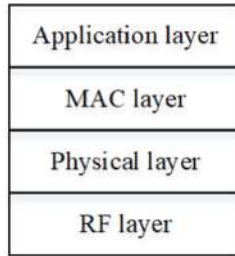


Fig. 3. LoRaWAN protocol layers

The RF layer defines the radio frequency bands that could be used in LoRa, namely for Europe and USA. We adopted the 868 MHz band available for Industry, Scientific and Medical (ISM) applications in Europe.

The LoRa physical layer implements a derivative of the Chirp Spread Spectrum (CSS) scheme. CSS was first developed in 1951 at Bell Telephone Laboratories for the military radars. It aimed to offer the same efficiency in range, resolution and speed of acquisition, but without the high peak power of the traditional short pulse mechanism. The MAC layer defines 3 classes of end nodes, respectively Class A, B and C. In this project we use only Class A, since it is the most energy efficient and the only one that is mandatory. To achieve this high energy efficiency, the nodes in this class are 99% of the time not active (neither transmitting nor receiving) and are only ready to receive immediately after transmitting a message. The Application layer is related with the user application layer.

As LoRa is a communication technology dealing with many connected nodes, it needs a robust end-to-end security. LoRa achieves this by implementing security at two different levels: the first one at the network level and a second one at the application level. Network security ensures authenticity of the node in the network and application security ensures that the operator has not access to the end user's application data.

The basic components of the LoRaWAN architecture are the following: nodes, gateways and network server. The nodes are the elements of the LoRa network where the sensing or control is undertaken. The gateway receives the data from the LoRa nodes and then transfers them into the backhaul system. The gateways are connected to the network server using standard IP connections. On this way the data communication uses a standard protocol, but any other communication network, either public or private, can be used. The LoRa network server manages the network, acting to eliminate duplicate packets, scheduling acknowledgements, and adapting data rates. The network server is also included in the CS, as it happened in the Wi-Fi solution. Figure 4 shows a simplified protocol stack of the three components interconnected.

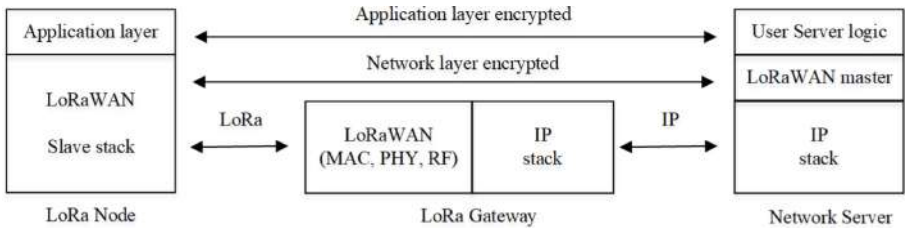


Fig. 4. Protocol stack of the LoRa based access network

The high level communication between the AdP meter and the CS (see Fig. 1) uses the Message Queuing Telemetry Transport (MQTT) protocol. MQTT is an IoT connectivity protocol. It was designed as a lightweight broker-based publish/subscribe messaging protocol, which is open, simple, lightweight and easy to implement. These characteristics make it ideal for use in constrained environments, for example, where the network has low bandwidth or is unreliable, as is the case of wireless sensor communications, or when run on an embedded device with limited processor or memory resources, as is the case of the AdP meters.

The AdP meter contains a MQTT client and the CS a MQTT Server or Broker. Periodically, e.g., every 5 min, the meter sends a MQTT Publish message to the MQTT Broker, located in the CS, with the following structure: Meter ID, Voltage, Current, Power, Power Factor, Energy, Service Time, Timestamp. For configuration of the different parameters in the AdP Meter the MQTT Server uses Subscribe messages with different topics, namely: Change of the measurement period, Change of the communication parameters (specific of Wi-Fi or LoRA), Set date/time, Set the initial value of the energy counter, Set current transformer ratio, Set Power Threshold (defines the power threshold to consider the equipment is in service), Set Meter mode (tri-phasic or 3 x mono-phasic).

In the MQTT architecture the elements that generate information are called Publishers and the elements that receive information are called Subscribers. The Publishers and Subscribers are interconnected through the MQTT Broker, as shown in Fig. 5.

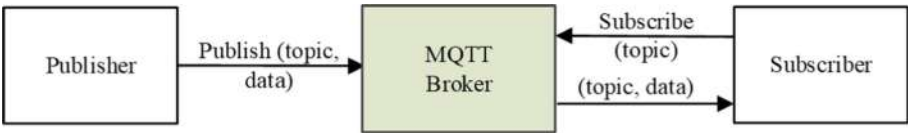


Fig. 5. MQTT communication architecture

The CS transmits the received information to the database by using the Structured Query Language (SQL). In the CS a software module converts the received messages coming from the Wi-Fi or LoRa based meter to a SQL message and transmits it to the database.

4 The AdP Meter

The AdP meter was designed to allow the monitoring, not only from energy consumption, but also of other electrical parameters like current, voltage, power and power factor. The AdP meter has the electrical interfaces shown in Fig. 6, and can be connected to a tri-phase electrical system.

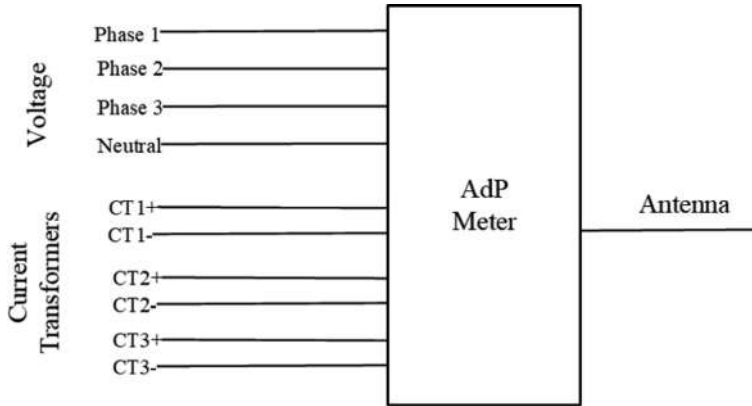


Fig. 6. AdP meter interfaces

There are 4 inputs for the three voltage phases and neutral connection on the top left. On the bottom left, there are 6 inputs for the connection of 3 current transformers, one for each phase. In Fig. 7, we can see the physical layout of the AdP meter, with the Voltage connectors and the antenna connector on the top and the current connectors on the bottom.



Fig. 7. AdP meter prototype

The AdP meter block diagram is shown in Fig. 8 and comprises four main modules: Measurement module, Processing and Communication module, Galvanic Isolation module and AC/DC dual power supply module.

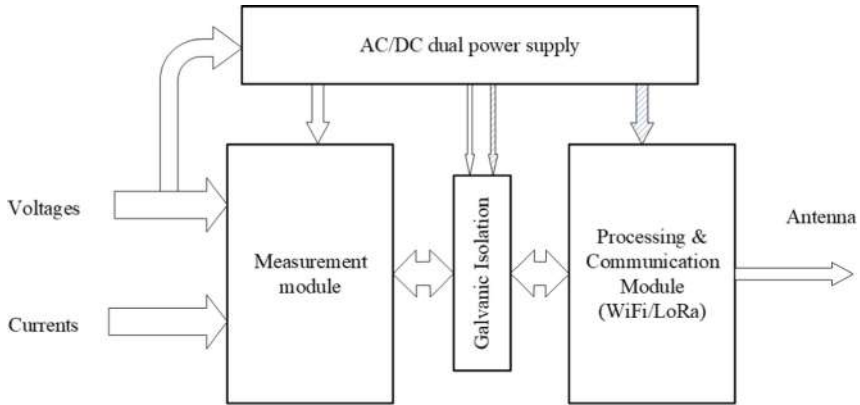


Fig. 8. AdP meter block diagram

The Measurement module and the Processing and Communication module are galvanic isolated for user protection, namely for the antenna connector and antenna cable handling. This requires a dual power supply with one of them connected to the Measurement module and the other to the Processing and Communication module. The Galvanic Isolation module requires to be connected to both.

The power, power factor, energy consumption and service time are calculated internally in the AdP meter from the voltage and current readings. They are transmitted to the CS, via MQTT protocol messages, using the following units for the different parameters: Voltage: 0.1 V, Current: 0.1 A, Power: Watt, Power Factor: 0–100 (100 correspond to $PF = 1$), Energy: 0.1 kWh (accumulated value), Service Time: minutes (accumulated value), Time Stamp: seconds.

5 Deployment

The meters are being tested in two large WRRF in the Lisbon area, named Chelas and Beirilas. Meters with the Wi-Fi module are installed in Chelas, while meters with the LoRa module are installed in Beirilas. The first objective is to test both communication technologies in order to make an evaluation of their strong and weak aspects, from the technical and economic points of view. The second objective is that AdTA is able to perform demand side management operations. By having the knowledge on the real-time energy consumption and on the values of other electrical parameters, in a demand management situation, AdTA will be able to shift loads in a controlled way so that the impact is minimized in the WRRF.

The Chelas WRRF covers an area of around 37,500 m² (250 m × 150 m) in a central area of Lisbon. Figure 9 shows the plant of the Chelas WRRF where Pi signals the points where the meters equipped with Wi-Fi are located. There are two meters in each Pi. The meters are connected to different WRRF equipment, such as pumps and ventilators. The total number of Wi-Fi APs to be deployed depends on the result of the on-going communication tests.

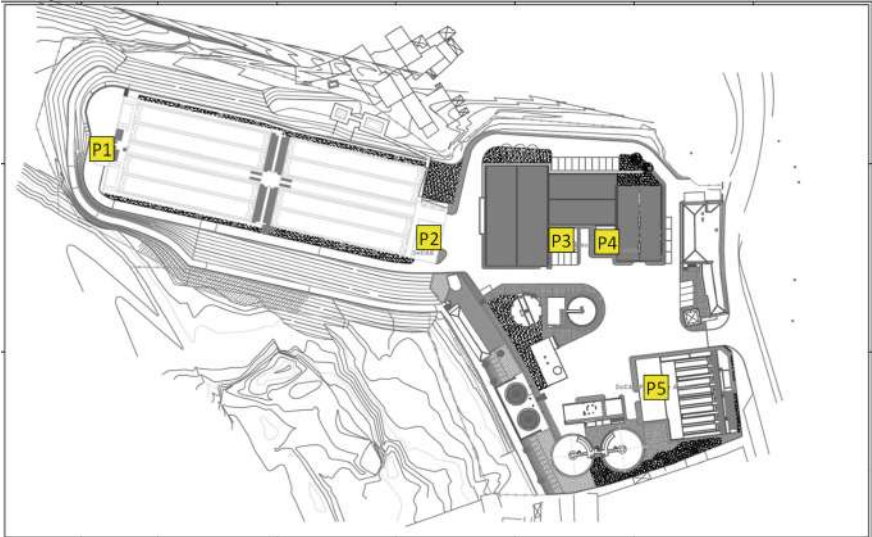


Fig. 9. Chelas WRRF

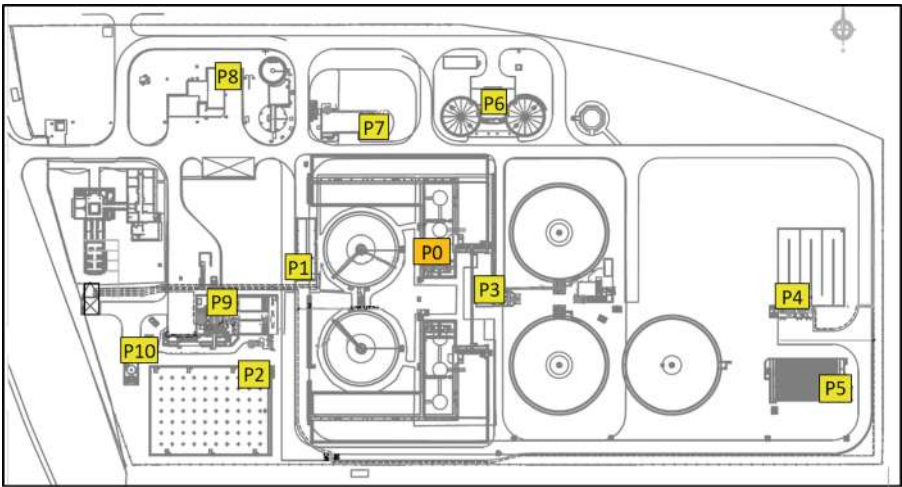


Fig. 10. Beirolos WRRF

The Beirolas WRRF covers a larger area of around 100,000 m² (400 m × 250 m) in Lisbon. Figure 10 shows the plant of the Beirolas WRRF where P_i are the points where the LoRa meters are located. There are also two meters located in each P_i . P_0 is the location of the LoRa gateway and antenna. The communication tests have already been performed and have validated this configuration with a single LoRa gateway.

6 Conclusion

An IoT based platform for real-time management of energy consumption in WRRF was presented. The developed work included the design and implementation of the adequate meters for measuring different electrical parameters (including energy consumption), the deployment of those meters in two WRRFs in the Lisbon area, the economical and performance analysis of two IoT communication protocols (Wi-Fi and LoRa) for access networks in the WRRF and the transmission of the data from either Wi-Fi APs or LoRa gateway to a central platform and database, where data analytics will be performed.

The objective of the pilots is, in first place, to decide on the communication technology to be used, when a more extended deployment of the system is done for other WRRF. The second objective is to be able to do demand side management operations, having in view the shifting of loads from peak load situations so that a better balance of the energy consumption can be achieved.

This work has to do with the so-called smart wastewater management, which is an important component of the smart city concept. It is worthwhile noticing too, that the project uses IoT technologies and architecture, which makes it up to date with the status of communications and platforms in smart cities. The future work includes the running of the platforms, extraction of the meter data and performance of data analytics on those data, so that guidelines can be designed for the extension of the platform to other WRRF.

Acknowledgment. The research leading to this work is being carried out as a part of the InteGrid project (Demonstration of INTElligent grid technologies for renewables INTEgration and INTERactive consumer participation enabling INTERoperable market solutions and INTERconnected stakeholders), which received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under grant agreement No. 731218. The sole responsibility for the content of this publication lies with the authors.

References

1. United States Environmental Protection Agency, Energy efficiency for water and wastewater utilities (2015). <https://www.epa.gov/sustainable-water-infrastructure/water-and-energy-efficiency-utilities-and-home>
2. International Water Association, Water Utility Pathways in a Circular Economy (2016). http://www.iwa-network.org/wp-content/uploads/2016/07/IWA_Circular_Economy_screen.pdf

3. AQUASAFE: an R&D complement to Bonn Network tools to support water safety plans implementation, exploitation and training, IWA Newsletter, vol. 1, no. 3 (2009)
4. Palensky, P., Dietrich, D.: Demand side management: demand response, intelligent energy systems, and smart loads. *IEEE Trans. Ind. Inform.* **7**(3), 381–388 (2011)
5. Aymerich, I., Rieger, L., Sobhani, D., Rosso, D., Corominas, L.: The difference between energy consumption and energy cost: modelling energy tariff structures for water resource recovery facilities. *Water Res.* **81**, 113–123 (2015). <https://www.sciencedirect.com/science/article/pii/S0043135415002705>
6. Póvoa, P., Oehmen, A., Inocêncio, P., Matos, J.S., Frazão, A.: Modelling energy costs for different operational strategies of a large water resource recovery facility. *Water Sci. Technol.* (2017). <https://doi.org/10.2166/wst.2017.089>
7. SmartWater4Energy project (2015). <http://smartwater4energy.hidromod.pt/>
8. Grilo, A., Casaca, A., Nunes, M., Bernardo, A., Rodrigues, P., Almeida, J.: A management system for low voltage grids. In: *Proceedings of the 12th IEEE PES PowerTech Conference, (PowerTech 2017)*, Manchester, United Kingdom, June 2017
9. Silva, N., et al.: Fault detection and location in low voltage grids based on distributed monitoring. In: *Proceedings of the IEEE International Energy Conference (ENERGYCON 2016) Conference*, Leuven, Belgium, April 2016
10. Keysight Technologies, *The Internet of Things: Enabling technologies and solutions for design and test* (2017). <https://literature.cdn.keysight.com/litweb/pdf/5992-1175EN.pdf?id=2666018>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

