Mehar Ullah

# FRAMEWORK FOR DIGITALIZING DIFFERENT INDUSTRIAL SECTORS VIA THE INTERNET OF THINGS

ACTA UNIVERSITATIS LAPPEENRANTAENSIS 1117

Mehar Ullah

# FRAMEWORK FOR DIGITALIZING DIFFERENT INDUSTRIAL SECTORS VIA THE INTERNET OF THINGS

Dissertation for the degree of Doctor of Science (Technology) to be presented with due permission for public examination and criticism in the Auditorium 1316 at Lappeenranta–Lahti University of Technology LUT, Lappeenranta, Finland on the 18th of December, 2023, at noon.

Acta Universitatis
Lappeenrantaensis 1117

Supervisors Associate Professor (Tenure Track) Pedro H. Juliano Nardelli
LUT School of Energy Systems
Lappeenranta–Lahti University of Technology LUT
Finland

Associate Professor (Tenure Track) Annika Wolff
LUT School of Engineering Sciences
Lappeenranta–Lahti University of Technology LUT
Finland

Dr. Arun Narayanan
LUT School of Energy Systems
Lappeenranta–Lahti University of Technology LUT
Finland

Reviewers Associate Professor Antonio Marcos Alberti
Instituto Nacional de Telecomunicações
Brazil

Associate Professor Ana Maria Diniz Moreira
Department of Software Engineering
Universidade Nova de Lisboa
Portugal

Opponent Associate Professor Antonio Marcos Alberti
Instituto Nacional de Telecomunicações
Brazil

ISBN 978-952-412-030-2
ISBN 978-952-412-031-9 (PDF)
ISSN 1456-4491 (Print)
ISSN 2814-5518 (Online)

Lappeenranta–Lahti University of Technology LUT
LUT University Press 2023

# Abstract

**Mehar Ullah**
**Framework for Digitalizing Different Industrial Sectors via the Internet of Things**
Lappeenranta 2023
66 pages
Acta Universitatis Lappeenrantaensis 1117
Diss. Lappeenranta–Lahti University of Technology
ISBN 978-952-412-030-2, ISBN 978-952-412-031-9 (PDF), ISSN 1456-4491 (Print), ISSN 2814-5518 (Online)

The Industrial Revolution 4.0 (IR4.0) refers to the current era of technological advancements that are transforming the different sectors of productive activities. It brings what is called Industry 4.0 characterized by the convergence of physical and digital systems that support the automation of several tasks. In this context, the Internet of Things (IoT) is enabling companies to make more informed decisions and operate more efficiently by providing a data network of physical objects that are connected to one another through dedicated applications. The industrial IoT has emerged as a key component of Industry 4.0, as it enables the collection of vast amounts of industrial data that can be used to optimize processes and improve decision-making. Data platforms have been developed to build and run different IoT applications, i.e., to process the huge amount of data generated by the IoT devices that are as diverse as the different industrial sectors. Handling such a massive amount of data is a challenge, leading to an approach usually called big data.

This doctoral dissertation explores different aspects of the IoT and explains how different parts of the IoT can work together to support a given application, especially in the industry. The focus is on the IoT data platforms—systems that enable the deployment and management of IoT devices. They are essential for data collection, analysis, and visualization by enabling a set of tools and services for device management, data analytics, and application development, as well as for support of various communication protocols and standards. Despite the generality of these processes related to the IoT, each industrial sector has specific requirements.

To enable companies to implement their specific IoT applications, an IoT platform is needed. The market offers a multitude of IoT platforms, each sharing similar functionalities but differing in their implementation and underlying technologies. These technological advancements present numerous difficulties for businesses and government entities, especially when dealing with the IoT infrastructure and platforms, which may be unfamiliar to many players in the field. Choosing an appropriate IoT platform from the available choices is a complex undertaking because this decision must consider not only current requirements but also potential future demands. This dissertation aims to answer the question: "Is it possible to create a unified framework for the digitalization of industrial sectors based on the IoT in integration with technologies like big data and analytics and edge computing?" The study also aims to answer the following subquestions: 1) How would a unified framework look like that can be used for the selection of an IoT platform based on companies' business requirements? 2) What happens when such a unified framework is applied to a specific domain of Industrial Energy Management? and 3) Can such an approach be deployed in different industrial sectors?

This dissertation offers a unified approach to solve practical deployment issues when digitalizing operations, taking into consideration particular applications. The main contributions of the dissertation can be summarized as follows. First, 21 key factors of an IoT platform required for the selection of a suitable IoT platform are identified for different applications considering the indications provided by the management of the industry, following a five-stage procedure. Second, a theoretical framework for an efficient cyber-physical system design is proposed by covering processes from data collection to end-user decision-making in order to build an industrial energy management system (IEnMS). Third, four different solutions based on the proposed approach are constructed for a diverse set of industrial applications, namely digitalization of a power-to-X plant, a cyber-physical pyrolysis process to recycle carbon fiber-reinforced polymer composite wastes, IoT platform selection for an IEnMS, and the data processing architecture of smart grids.

**Keywords:** Industrial IoT, IoT platforms, industrial energy management system, cyber-physical systems

## Acknowledgments

This study was carried out at the LUT School of Energy Systems at Lappeenranta–Lahti University of Technology LUT, Finland, between 2019 and 2023. These four years spent at LUT University were more than an excellent journey for me. I have had the chance to meet and add new people to my life. I have met new students, professors, professionals in companies, and cricket players, among others.

I would like to extend my thanks to all who have guided, supported, encouraged, and advised me throughout my doctoral studies here at LUT during my course work and also in my dissertation time. First of all, I would like to express my thanks to my supervisors Associate Professor (Tenure Track) Pedro Nardelli and Associate Professor (Tenure Track) Annika Wolff. Their guidance and support made me able to complete my courses, research papers, and dissertation tasks. Pedro Nardelli and Annika Wolff have motivated me throughout my studies. Pedro is such a nice and helpful professor, supervisor, and friend who has provided me a comfortable environment in the office, polished my research work, and continuously helped me throughout my studies.

I would like to say thanks to Professor Kari Smolander for his support in my initial days of starting research at LUT University. My LUT journey started just because of him. Special thanks to all my Cyber Physical System group members including Mohammad, Majid Hussain, Daniel, Arun, Dick, Ali, Pedro Goria, Daynara, Arthur, and Cassia for providing such a nice and friendly environment throughout the study time. I am also so much thankful to all the staff of Lappeenranta–Lahti University of Technology for providing me the facilities and services during my studies.
This work was possible because of the funding received from the Academy of Finland (now the Research Council of Finland) through the projects FIREMAN (n. 326270 as part of the grant CHIST-ERA-17-BDSI-003), and EnergyNet (n. 328869), and from Business Finland through the HYGECEL project. In addition, I would like to thank the support from the MORE SIM LUT research platform and the LUT Research Foundation.

Moreover, I would like to say thanks to my Mom and Dad for their prayers, thanks to which I have been able to reach this position. My elder brother Adnan Ahmad is a strong wall behind my back; he has also supported me in my life, and I can never imagine a happy life without him.

Finally, I would like to dedicate my dissertation to my sweet wife Kiran and my sons Hamdan and Hayyan, who have made my life so beautiful and prosperous by supporting me throughout my studies.

Mehar Ullah
December 2023
Lappeenranta, Finland

*To my sweet wife Kiran*
*and Angels Hamdan and Hayyan*

*Yours, Mehar*

# Contents

# List of publications

**Publication I**

Ullah, M., Nardelli, P. H. J., Wolff, A., and Smolander, K. (2020). "Twenty-One Key Factors to Choose an IoT Platform: Theoretical Framework and Its Applications." *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10111–10119, Oct. 2020, doi: 10.1109/JIOT.2020.3000056.

**Publication II**

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P. (2022). "Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data." *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3215167.

**Publication III**

Ullah, M., Gutierrez-Rojas, D., Inkeri, E., Tynjälä, T., and Nardelli P. H. J. (2022). "Operation of Power-to-X-Related Processes Based on Advanced Data-Driven Methods: A Comprehensive Review." *Energies* 2022, 15, 8118. https://doi.org/10.3390/en15218118

**Publication IV**

Ullah, M., Gopalraj, S. K., Gutierrez-Rojas, D., Nardelli, P., and Kärki, T. (2023). "IoT framework and requirement for intelligent industrial pyrolysis process to recycle CFRP composite wastes: application study." In: *Intelligent and Transformative Production in Pandemic Times: Proceedings of the 26th International Conference on Production Research*, pp. 275–282. Cham: Springer International Publishing, https://doi.org/10.1007/978-3-031-18641-7_26

**Publication V**

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P. H. J. (2021). "Unified Framework to Select an IoT Platform for Industrial Energy Management Systems." In: *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 950–955, doi: 10.23919/MIPRO52101.2021.9597128.

**Publication VI**

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P. (2021). "Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing." In: *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 956–961, doi: 10.23919/MIPRO52101.2021.9596885.

# List of other publications

The publications listed below are not included in the current dissertation, but have been produced in the course of the doctoral studies. In these publications, the author is either the main researcher or a coauthor of the publications.

### Journal publications

Narayanan, A., De Sena, A. S., Gutierrez-Rojas, D., Melgarejo, D. C., Hussain, H. M., **Ullah, M.**, Bayhan, S., and Nardelli, P. H. (2020). "Key advances in pervasive edge computing for industrial internet of things in 5g and beyond." *IEEE Access*, 8, pp. 206734–206754.

Narayanan, A., Korium, M., Melgarejo, D. C., Hussain, H. M., De Sena, A. S., Goria, P., Gutierrez-Rojas, D., **Ullah, M.**, Esmaeelnezhad, A., Rasti, M., and Pournaras, E. (2022). "Collective intelligence using 5G: Concepts, applications, and challenges in sociotechnical environments." *IEEE Access*.

Carrillo, D., Duc Nguyen, L., Nardelli, P. H., Pournaras, E., Morita, P., Rodríguez, D. Z., Dzaferagic, M., Siljak, H., Jung, A., **Ullah, M.**, Hebert-Dufresne, L., and Macaluso, I. (2021). "Containing future epidemics with trustworthy federated systems for ubiquitous warning and response." *Frontiers in Communications and Networks*, p. 11.

### Conference publications

**Ullah, M.** and Smolander, K. (2019). "Highlighting the key factors of an IoT platform." In: *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 901–906, IEEE.

Gutierrez-Rojas, D., **Ullah, M.**, Christou, I. T., Almeida, G., Nardelli, P., Carrillo, D., Sant'Ana, J. M., Alves, H., Dzaferagic, M., Chiumento, A., and Kalalas, C. (2020). "Three-layer approach to detect anomalies in industrial environments based on machine learning." In: *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, vol. 1, pp. 250–256, IEEE.

**Ullah, M.**, Hekmatmanesh, A., Savchenko, D., Moioli, R., Nardelli, P., Handroos, H., and Wu, H. (2020). "Providing facilities in health care via brain-computer interface and Internet of Things." In: *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 971–976, IEEE.

**Ullah, M.**, Narayanan, A., Wolff, A., and Nardelli, P. H. (2021). "Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data." In: *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1–6, IEEE.

**Ullah, M.**, Kakakhel, S. R. U., Westerlund, T., Wolff, A., Carrillo, D., Plosila, J., and Nardelli, P. H. (2020). "Iot protocol selection for smart grid applications: Merging qualitative and quantitative metrics." In: *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 993–998, IEEE.

**Book chapters**

**Ullah, M.**, Wolff, A., and Nardelli, P. H. J. (2021). "Processing Smart Meter Data Using IoT, Edge Computing, and Big Data Analytics." In: Fathi, M., Zio, E., Pardalos P. M. (Eds.). *Handbook of Smart Energy Systems*. Cham: Springer. https://doi.org/10.1007/978-3-030-72322-4-124-1.

Sena, A.S.d., **Ullah, M.**, and Nardelli, P. H. J. (2021). "Edge Computing in Smart Grids." In: Fathi, M., Zio, E., Pardalos, P. M. (Eds.). *Handbook of Smart Energy Systems*. Cham: Springer. https://doi.org/10.1007/978-3-030-72322-4-106-1.

# Nomenclature

### Abbreviations

| | |
|---|---|
| AI | Artificial intelligence |
| AMQP | Advanced message queuing protocol |
| API | Application programming interface |
| CFRP | Carbon fiber-reinforced polymer |
| CFs | Carbon fibers |
| $CO_2$ | Carbon dioxide |
| CoAP | Constrained application protocol |
| EnM | Energy management |
| EnMS | Energy management system |
| EU | European Union |
| GCP | Google cloud platform |
| $H_2$ | Hydrogen |
| HDFS | Hadoop distributed file system |
| HEMS | Home energy management system |
| HTTP | Hypertext transfer protocol |
| IaaS | Infrastructure as a service |
| ICT | Information and communication technology |
| IEnMS | Industrial energy management system |
| IIoT | Industrial Internet of Things |
| IoT | Internet of Things |
| IR4.0 | Industrial Revolution 4.0 |
| ML | Machine learning |
| mMTC | Massive machine type communication |
| MQTT | Message queuing telemetry transport |
| NSF | National science foundation |
| P2G | Power-to-gas |
| P2X | Power-to-X |
| PaaS | Platform as a service |
| RFID | Radio frequency identification system |
| Saas | Software as a service |
| SAMOA | Scalable advanced massive online analysis |
| SG | Smart grid |
| SNG | Synthetic natural gas |
| SQL | Sequential query language |
| US | United States |
| YARN | Yet another resource negotiator |

# 1 Introduction

The Industrial Revolution 4.0 (IR4.0), which includes cyber-physical systems (CPSs), industrial Internet of Things (IoT), edge computing, and artificial intelligence, is essentially a trend toward automation and data exchange in technologies related to different sectors of material production. Although there is currently no evidence that the expertise of humans will be replaced by machines, they are often more effective than humans in performing repetitive tasks, and the combination of machine learning and computational power enables machines to perform extremely difficult tasks [1].

To this end, a development called Industry 4.0, which will be more and more dependent on the IoT, is currently evolving. The use of digital technologies in industrial processes produces an ecosystem that is highly connected and intelligent. This technological evolution has a positive effect on diverse applications like industrial automation, predictive maintenance, energy management, and data-driven decision-making. Industry 4.0 can be characterized by interconnected, intelligent, and adaptable industrial systems utilizing the power of the IoT, which may then boost productivity, efficiency, and competitiveness in the industrial sector. Thus, expertise in the IoT and its supporting platforms in order to deploy different applications becomes a necessity.

Despite (or because of) the rapid technological development in information and communication technologies (ICTs), there are still fundamental questions to be solved, especially when actual solutions must be deployed and successfully operated. One of the problems is related to the selection of the most suitable choice for running the data processes, considering the large number of service providers and their specific architectures for data networks (also with the possibility of in-house operation of the industrial data network). Other problems arise from the mismatch between the demands of the representatives of a given industry and what the ICTs can offer.

## 1.1 Main objective and research questions

In this context, the main objective of this doctoral research is to propose a general, unified approach to support the selection and deployment of IoT-based solutions in different industrial sectors, considering both the specialized technical knowledge and opinions provided by experts and members of industrial organizations. Specifically, the proposed framework involves not only the IoT but also big data and analytics to effectively improve different applications in different industrial sectors. The main objective can be expressed as the following question: **Is it possible to create a unified framework for the digitalization of industrial sectors based on the IoT in integration with technologies like big data and analytics and edge computing?** In this dissertation, the top five IoT platforms (AWS, Microsoft Azure, Google cloud IoT, IBM Watson IoT, and Oracle IoT) have been selected based on their market shares for considering the key factors of an IoT platform.

To achieve this broad aim, three different intermediate research questions must be answered:

**RQ1:** What would a unified framework look like that can be used for the selection of an IoT platform based on companies' business requirements?

**RQ2:** What happens when such a unified framework is applied to a specific domain of Industrial Energy Management?

**RQ3:** Can such an approach be deployed in different industrial sectors?

A schematic view of the aforementioned research questions is given in Figure 1.1.

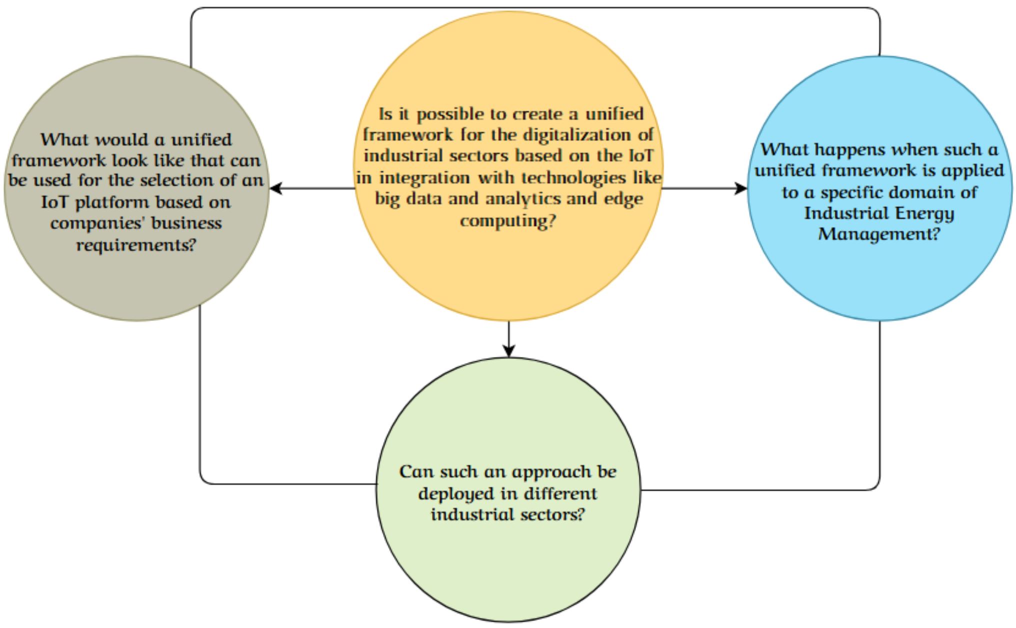



Figure 1.1: Schematic presentation of the research questions to be answered in this doctoral dissertation.

## 1.2 Hypothesis and contributions

Those RQs were posed based on the hypothesis that digitalization of industrial processes can be made effective through the IoT in connection with the modern technologies like big data and analytics and edge computing, leading to their cyber-physical operation. However, to achieve its goals, a proper selection of an IoT platform is needed considering the specifics of the end application; this was the aim of Publication I, which offers the answer to RQ1. Besides, to advance toward a specific case, which refers to RQ2, the attention was turned to an IoT-based deployment of Industrial Energy Management Systems (IEnMS) as a timely case study considering the current sustainability actions. This contribution indicated the path to implement an IEnMS as a technical tool that is part of the energy policies of a given company. The results were presented in Publication II and Publication V. The methodology of these two publications (I & II) is based on an in-depth literature review and expert opinions obtained through questionnaires.

To assess the generality of the proposed framework to select an IoT platform and design a specific data network architecture to acquire, process, and analyze data, four illustrative cases of digitalization of industrial operations were investigated, namely cyber-physical operation of a power-to-X plant (Publication III) and a pyrolysis process (Publication IV), IoT platform selection for an IEnMS (Publication V), and big data processing architecture for smart grids (Publication VI) as shown in Figure 1.2. These four contributions demonstrate the potential that the proposed unified framework has, answering RQ2 and RQ3. It is pointed out that the methodology employed in Publications III–VI was based on the utilization of the framework developed in Publications I and II but applied in different sectors. The research was conducted in collaboration with experts in their respective fields, bringing an interesting interdisciplinary aspect to these studies.

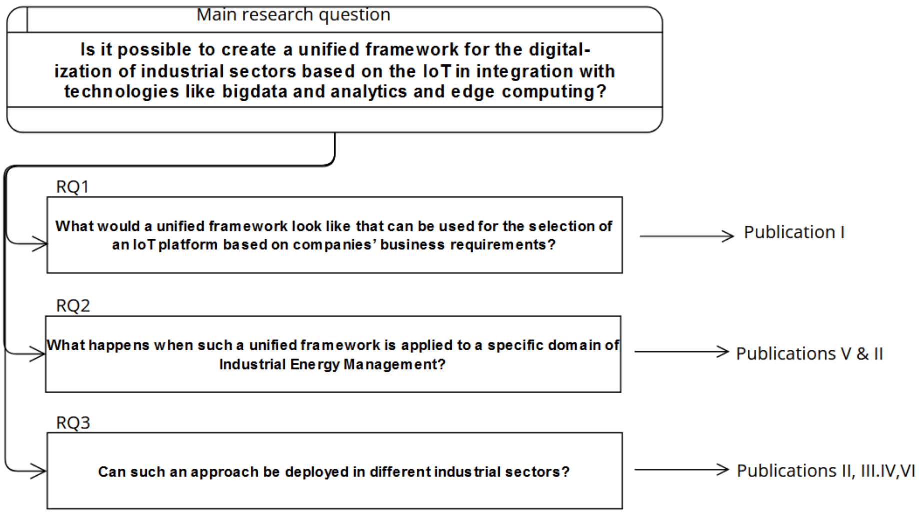



Figure 1.2: Schematic presentation of the research questions and publications answering the questions.

## 1.3 Structure of the dissertation

This doctoral dissertation is structured to answer the above-mentioned research questions in a sequential manner in the following chapters.

**Chapter 1** covers a brief introduction of all the topics included in the dissertation. The chapter also describes the motivation of the doctoral research and introduces the structure of the dissertation.

**Chapter 2** contains the background information about the IR4.0, the IoT, IoT platforms, big data, edge computing, and cyber-physical systems.

**Chapter 3** provides the results of the publications and answers the above-mentioned questions. The chapter highlights the contributions of Publications I–VI, which mainly focus on the proposed general framework for the digitalization of the operations of different industrial sectors considering their specific requirements and design options offered by the IoT platforms.

- Section 3.1 presents the results of Publication I, which answers RQ1. In the publication, the key factors of the IoT platform are identified from the literature, and five top IoT platforms are studied based on their market shares. Based on companies' business requirements, an IoT platform selection framework is designed.
- Section 3.2 highlights the contributions of Publication II, which provides information about the Industrial Energy Management System (IEnMS) and how the IoT and big data analytics are used to facilitate the IEnMS and its operations. This section answers RQ3.
- Section 3.3 contains the contributions and results of Publications III–VI, the focus being on the importance and role of the IoT and big data in various industrial applications including the energy sector (smart grids), an industrial pyrolysis process, and operation of power-to-X processes. This section answers RQ2 and RQ3.

**Chapter 4** concludes the dissertation with a summary of the main scientific contributions of the work, as well as its implications, limitations, and future perspectives.

# 2 Background

## 2.1 Internet of Things

The Internet of Things, or simply the IoT, refers to how physical objects are connected to one another over a data network. The IoT uses sophisticated backend systems that require various capabilities based on the system's needs. The IoT is already used in many realms of daily life, including agriculture, homes, health services, transport, and power grids [2]. The idea of the IoT was first introduced by Kevin Ashton in 1999 [3]. After almost 25 years, the IoT is still gaining traction in a number of industrial sectors including logistics, industrial manufacturing, and multi-energy systems. Connecting the physical and digital worlds is the main goal of the IoT, which initially used radio frequency identification system (RFID) technology to identify, track, and monitor different physical elements. IoT applications today incorporate various types of data acquired from sensors and further processed in computing devices, moving further and further away from basic RFIDs [4].

### 2.1.1 Key components of the IoT

The functionalities and importance of the IoT, which are identified in the literature review presented in Publication I, can be described by addressing its building blocks. There are six building blocks, as presented in Figure 2.1. These blocks are the components of the IoT that are working together to support specific applications and solve specific tasks [2].

**Identification block:** There is a high integration of technologies in the IR4.0, resulting in a great amount of digitalization and networking in industries, and the number of devices and the communication involved have increased [5]. The identification block is used for the identification and tracking of devices within the huge number of devices/objects in the IoT network to manage and control these devices in an efficient and secure manner. Devices in the network are identified by their object ID, which contains their name and address, by using different protocols, such as IPv6 and IPv4, and different methods like electronic product codes (EPC) and ubiquitous codes (uCode). Because the identification techniques are not universally unique, it is crucial to distinguish between object identification and address [2]. Addressing helps to uniquely identify objects. The identification block plays a major role in the IR4.0 as it acts as the foundation for the connectivity, interoperability, security, and data-driven decision-making, which are required for the industrial automation.

**Sensing block:** Sensors acquire data of specific processes to then send them to a computing unit (usually in the cloud) that can handle their processing. Often, there are also actuators or physical hardware like switches that intervene in the physical world through commands generated in IoT platforms, working in opposition to sensors [6]. This is the basis of control systems and cyber-physical systems [7].

**Communication block:** The connectivity part of the IoT is one of the main elements of

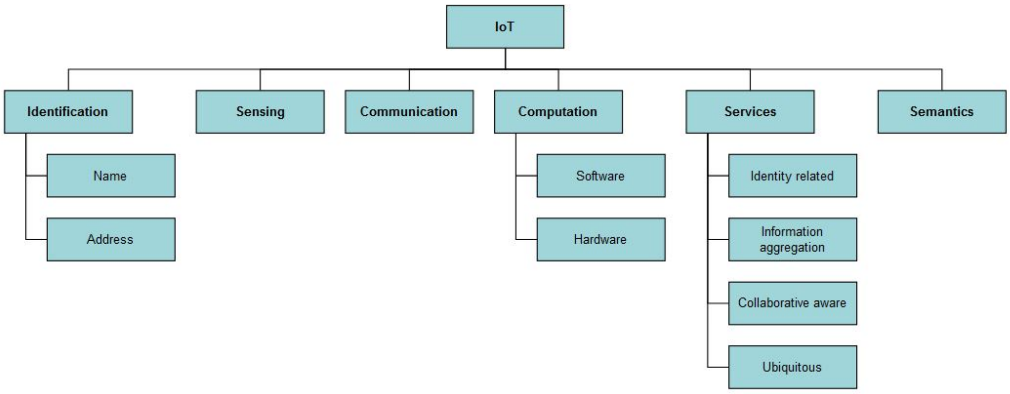



Figure 2.1: Building blocks of the IoT (adapted from Publication I).

the IR4.0 as it provides the physical systems with digital technologies, resulting in the automation of industrial processes. Connectivity enables collection of data from sensors and elements (things), real-time decision-making, and integration of digital technologies into industrial processes. In the IR4.0, the selection of connectivity technologies depends on several factors, such as data rate, range, application requirements, and reliability. In this process, many heterogeneous objects communicate with each other and with the platform to exchange information related to different services by using protocols like MQTT and CoAP. Through the use of communication technologies like ZigBee, Near Field Communication (NFC), wireless fidelity (Wi-Fi), SigFox, Long Range (LoRa), and cellular networks, sensors and other devices are linked to the Internet [2, 8].

**Computation block:** The computation block of the IoT plays a major role in the IR4.0 as it provides the required computing power, performs data processing, providing intelligence, and connects different industrial processes, thus bringing benefits to the IR4.0 [9]. There are two components in the computation block: hardware and software. IoT applications can be run on many hardware platforms, such as Intel Galileo, Raspberry PI, UDOO, and Arduino. Like hardware platforms, software platforms are used to carry out IoT functionalities. The operating system, which is used almost exclusively during the device activation process, is the primary software platform [2, 8].

**Services block:** This block comprises the environment that the developers use when creating new applications considering the existence of already available services, which can be classified into four categories. There are two types of identity-related services: active and passive. Active identity-related services are those that broadcast information and use battery power or a constant power source. Information can be sent or transmitted to another device by active identity-related services. Passive identity-related services require an external device or mechanism to transmit their identity, because they lack a power source. Services that deal with passive identities can only read data from objects. The processes of gathering data from sensors, processing those data, and then sending them to the IoT application for processing are referred to as information aggregation services. The information provided by the information aggregation services is used by collaborative-aware services to make decisions and take appropriate actions. Everybody who needs them can get collaboratively aware services from ubiquitous services at any time and anywhere [10, 11].

**Semantics block:** To efficiently acquire this knowledge, contextualized data processing leading to cognitive features of the data based on information is required. In this case, objects are associated with a semantics that is based on, for example, location and utilization of a resource, its model, and predicted behavior [2, 8].

## 2.2 IoT platforms

IoT platforms are deployed to provide services to applications that are related to management, connectivity, and data processing, as well as cyber-security and user interfacing [12]. As reported in Publication I, there has been a rapid growth in the number of IoT platforms; in 2020, there were at least 620 IoT platforms, leading to a fierce competition between service providers [13]. On the other hand, as reported in 2018, the majority of users are served by the five largest IoT platforms, namely Amazon Web Services (AWS), Google Cloud, Microsoft Azure, IBM Watson IoT, and Oracle. This, however, brings challenges for the management and decision-makers of different companies if they wish to digitalize their internal processes through the IoT.

For companies to select an appropriate IoT platform, a few important factors must be taken into account based on the specific needs of the application under consideration. It is not necessary for an IoT platform to have all possible features; rather, the decision should be a balance between the features for a well-determined context. By examining different IoT platforms, these factors were identified in the literature [8, 14, 15] and reported in Publication I.

In this dissertation, the key elements of the IR4.0 are explained [5, 9, 16, 17, 18, 19, 20, 21]. Keeping these points in mind, it is discussed in the following how the IoT platform selection should be carried out.

**Stability:** In terms of the IR4.0, stability refers to the ability of businesses, industries, and societies to maintain a sense of equilibrium, security, and predictability in the face of rapid technological advancements and disruptions. While the IR4.0 offers numerous opportunities for growth and innovation, it also presents challenges that need to be addressed to ensure stability. There may be some platforms that are unable to provide customers with such services. Therefore, a platform that has a good chance of surviving in the market should be chosen. Here, customers who have previously used the platform can provide information about it [8].

**Scalability and flexibility:** In the IR4.0, the term scalability refers to the ability of businesses and industries to adapt, grow, and evolve in response to the rapid advancements in technology and changing market demands. The IR4.0 represents a new era of industrialization characterized by the integration of digital technologies, automation, data analytics, and the IoT into various aspects of manufacturing and business processes. Initially, a company may be small, but it may grow over time. Therefore, the IoT platform should be scalable to the potential growth of the company. The platform should also be adaptable in terms of technology, because both the consumer demand and modern technologies are constantly changing [22].

**Pricing model:** The pricing model for the Internet of Things (IoT) plays a significant role in the context of the IR4.0. The IR4.0 is characterized by the integration of digital technologies, including the IoT, into various aspects of industrial and manufacturing processes. The IoT pricing models need to be flexible and scalable to accommodate the unique needs of businesses in this era. At the beginning of a contract agreement, some platform providers offer a low price; however, after that, the price may rise sharply. Additionally, some service providers advertise low prices to attract clients, but the contracts contain only a few features, and adding more features would be expensive. Consequently, such a platform should be chosen that provides all the features that the company needs at a price that is in line with its budget [23].

**Security:** In the context of the IR4.0, security is a critical concern because of the increasing integration of digital technologies, automation, and connectivity in industrial processes. Ensuring the security of systems, data, and operations is essential to prevent cyberattacks, protect sensitive information, and maintain the reliability and safety of critical infrastructure. Security is a crucial component of the IoT, the quality of which should be high in all platforms. Data encryption, application authentication, secure session initiation, application authentication, cloud security, and device security are just a few possible types of security that can be used [24].

**Time to market:** The IR4.0 is characterized by the integration of digital technologies, data-driven decision-making, and automation into various aspects of manufacturing and business processes. Reducing time to market is a key objective in the IR4.0 because it can lead to numerous advantages, including competitiveness, innovation, and responsiveness to changing customer demands. The questions of time to market and how the platform provider will support the company during the process from product conception to sale should be taken into account when choosing an IoT platform. Some IoT platform providers provide new clients with quick start packages that can expedite product development, shorten the time to market, and provide better solutions [24].

**Data analytics and visualization tool:** Data analytics and visualization are fundamental components of the IR4.0, enabling organizations to extract useful insights from the vast amounts of data generated by connected devices and processes. These insights drive improvements in efficiency, productivity, quality, and innovation, ultimately contributing to the success of the IR4.0. Prospective IoT platform users should determine which platform offers the best capabilities to aggregate, analyze, and visualize data before choosing an IoT platform. Users should pay close attention to how the IoT platform replaces built-in functionalities by integrating top analytics tools. Prior to choosing an IoT platform, requirements for data analysis and information visualization should be determined [24].

**Data ownership:** One of the complex and evolving issues of the IR4.0 is the data ownership. The IR4.0 is characterized by the widespread use of digital technologies, the Internet of Things (IoT), artificial intelligence (AI), and big data analytics, all of which generate and rely on vast amounts of data. The ownership of these data is a critical concern, which raises legal, ethical, and economic questions. With IoT data, the ownership of the data is a challenging issue. Further, laws vary between jurisdictions. For instance, the European Union (EU) and the United States (US) have different laws and policies regarding data ownership. The service provider of an IoT platform then must have knowledge of data rights and the geographic scope of data protection [24].

**Cloud infrastructure ownership:** The ownership of the cloud infrastructure in the IR4.0 is a complex and evolving landscape. The IR4.0, marked by the integration of digital technologies, automation, and data-driven processes, relies heavily on cloud computing for storage, processing, and data accessibility. The public cloud infrastructure is owned and operated by third-party cloud service providers, such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and others. These companies own and manage vast data centers and offer cloud services on a pay-as-you-go basis. Many businesses and organizations use public cloud services to host their applications, data, and services, effectively outsourcing the infrastructure ownership to these providers. Some smaller IoT platform providers only offer the software layer because the hardware infrastructure layer is expensive. Some service providers run their services primarily on a single top platform and certify their platform on one or more of the top public cloud service providers. It is important to verify that the provider of the IoT platform is compatible with the larger enterprise cloud [22].

**Extension of legacy architecture:** The extent of legacy architecture in the IR4.0 varies based on industry, organization, and various factors, such as cost, regulation, and risk tolerance. While many businesses are working to embrace digital transformation and replace legacy systems with more advanced technologies, the process can be complex and time-consuming, leading some organizations to continue using legacy systems alongside newer innovations. IoT devices are made to function with a variety of infrastructure systems, but the connectivity in an existing IoT is frequently unknown. Therefore, organizations should consider how new technology generations can work together with older technology when choosing an IoT platform.

**Protocols:** In the IR4.0, protocols refer to standardized sets of rules and conventions that enable seamless communication, interoperability, and data exchange between various devices, systems, and components in a connected and automated environment. These protocols play a crucial role in ensuring the efficient operation of IR4.0 technologies.

MQTT, HTTP, AMQP, and CoAP are some of the crucial protocols that are supported by IoT platforms. For example, MQTT is very light and has much lower overheads because it is binary. As new IoT-enabled devices are constantly being introduced, the chosen IoT platform should be flexible to support the inclusion of new protocols that are brought by such a development.

**System performance:** In the IR4.0, system performance is a critical aspect that pertains to the efficiency, effectiveness, and reliability of various technological systems and processes used in modern manufacturing and business operations. In the IR4.0, where technologies like automation, data analytics, the IoT, AI, and advanced robotics are heavily integrated, optimizing the system performance is essential for achieving the desired outcomes. When an event occurs in an IoT platform, a rule-based trigger may be automatically activated. The average time to analyze and handle each event increases as more devices connect to the IoT platform because they support such a method. Prior to choosing an IoT platform, it is important to understand the measures the provider has taken to keep the platform performance at a high level.

**Interoperability:** In the IR4.0, interoperability is a critical concept that relates to the ability of different systems, devices, machines, and software to connect, communicate, and work together in a seamless manner. This interoperability is fundamental to the success of the IR4.0 as it enables the integration of various technologies and components, facilitating data exchange and automation. Middleware refers to the software developed to be in the middle between the end application and the servers supporting the IoT platform. Numerous end applications will use the collected data, and the platform itself may not even have access to it. The chosen IoT platform should thus enable integration with open-source ecosystems. The organization that is using the IoT platform can increase its productivity by ensuring interoperability of the middleware along different data sources.

**Redundancy and disaster recovery:** In the IR4.0, where downtime and data loss can have significant financial and operational consequences, redundancy and disaster recovery are critical investments. These strategies help organizations maintain operational stability, protect valuable data, and mitigate the impact of disruptions, whether they are caused by hardware failures, cyberattacks, natural disasters, or other unforeseen events. IoT platform providers should have specialized infrastructure to handle data during adverse events (e.g., disasters). Problems related to the IT infrastructure can occasionally occur, either caused by natural disasters or human errors. The data backup plan schedule and the availability of failover cluster provision in the IoT platform are issues that must be taken into account before the selection of an IoT platform.

**Attractive interface:** In the IR4.0, an attractive interface refers to user interfaces (UI) and user experiences (UX) designed to be visually appealing, intuitive, and efficient for users interacting with digital technologies, automation systems, and smart devices. To create an attractive interface in the IR4.0, businesses often collaborate with UI/UX designers and usability experts who specialize in designing interfaces that are visually appealing, user-friendly, and aligned with the specific needs of the industry. In order to make it simple for users to access the platform features, the interface of the IoT platform has to be appealing, user-friendly, and straightforward. All of the services provided to the customers should be simple to access.

**Application environment:** The business application environment in the IR4.0 is characterized by the integration of advanced digital technologies and automation into various aspects of business operations. This transformation has a profound impact on how businesses operate, innovate, and compete in the global marketplace. The IR4.0 involves digitalization of the core business processes. This includes the use of sensors and IoT devices to collect real-time data, cloud computing for data storage and processing, and adoption of digital workflows to streamline operations. Before choosing an IoT platform, it is important to take into account three aspects of the application environment: the types of applications that are preinstalled, the features of the application development environment, and the common interfaces.

**Hybrid cloud:** Hybrid cloud solutions play a significant role in the IR4.0 owing to their ability to address the diverse and evolving needs of businesses in the digital era. IR4.0 technologies generate large volumes of data, which may require a flexible and scalable IT infrastructure. A hybrid cloud allows businesses to scale their computing and storage resources up or down based on demand, ensuring that they can handle the data-intensive workloads of the IR4.0 without significant upfront investments. Hybrid cloud solutions are essential for businesses in the IR4.0, because they provide the agility, scalability, and flexibility required to leverage the full potential of digital transformation while effectively managing data, legacy systems, security, and compliance. They serve as a bridge between the traditional infrastructure and the cloud-native, data-driven future of the IR4.0. Some IoT platforms can integrate with current IT systems that are hosted on the premises of the company. In these circumstances, a hybrid cloud is very helpful because it allows the management of public and less important operations by the platform while mission-critical or sensitive processes can be handled locally.

**Platform migration:** Platform migration in the context of the IR4.0 is the process of transitioning from legacy systems or platforms to new, digital, and often cloud-based platforms that leverage advanced technologies to drive innovation, efficiency, and competitiveness. Platform migration is a crucial strategic move for many organizations as they seek to adapt to the changing industrial landscape. Before migrating to a new platform, organizations need to conduct a comprehensive assessment of their current systems, processes, and technology infrastructure. This assessment should identify the shortcomings and opportunities for improvement that necessitate a migration. The IoT platform might not be able to accommodate all the needs of the company over time as it expands. Consequently, a larger IoT platform provider might be required. In order to prepare for any potential future migration to other IoT platforms, companies should ensure that the chosen IoT platform provider offers clearly documented interfaces, schemas, and application programming interfaces (APIs).

**Previous experience:** Before making a decision, a company should determine whether the IoT platform provider has experience with the work of the companies that are following the IR4.0, and whether this experience is comparable with the company application. A successful work history in the same field can be viewed favorably.

**Bandwidth:** In the IR4.0, bandwidth plays a crucial role in enabling the seamless flow of data between devices, machines, sensors, and systems, which are interconnected as part of the Industrial Internet of Things (IIoT). It allows the rapid and reliable exchange of data, supports real-time decision-making, and underpins the interconnected nature of smart factories and industrial processes in the IR4.0. The availability and capacity of bandwidth will continue to be a key consideration for organizations looking to fully harness the benefits of the IR4.0. The IoT platform requires low-latency and high-bandwidth networking for effective data transfer and communication between the processing elements. Therefore, it is important to confirm that a potential IoT platform provider has a large data pipe and enough room for expansion.

**Edge intelligence and control:** Edge computing involves the deployment of advanced computing capabilities, data analytics, and control systems closer to the edge of the network, where data are generated and actions are executed. Edge intelligence and control are fundamental components of the IR4.0, enabling faster decision-making, greater efficiency, improved safety, and more responsive industrial processes. By processing data closer to the source and leveraging advanced technologies like AI and the IoT, businesses can gain a competitive edge and better adapt to the demands of the IR4.0. Decentralization of centralized data processing as in large-scale computing servers is also an option in some specific cases. This brings the computation close to the network edge. As this paradigm is becoming more frequent, it is important to ensure that the IoT platform can support novel topologies to make use of what is called edge intelligence.

## 2.3 Big data and analytics

The IR4.0 mostly requires the adoption of big data and the related technologies that are combined to meet the data collection, storage, processing, and analysis requirements [25]. Because of industrial automation in the IR4.0, a huge amount of data is generated by a variety of sources, including machine controllers, sensors, industrial equipment, people, social media, and software applications [26], and conventional data processing software and techniques cannot usually successfully process such big data in order to produce useful information [27]. The term big data refers to all of these massive data entering at rapid speeds and in various formats (structured, unstructured, and semistructured) [25]. "The three Vs": *volume, velocity*, and *variety*, are well-known big data properties. The amount of data generated is associated with volume, the speed with which data are processed is associated with velocity, and the variation in the sorts (types) of data produced is associated with variety [28]. Utilizing massive amounts of data in order to find meaningful insights, trends, or models is the key to long-term innovation in an IR4.0 factory. For

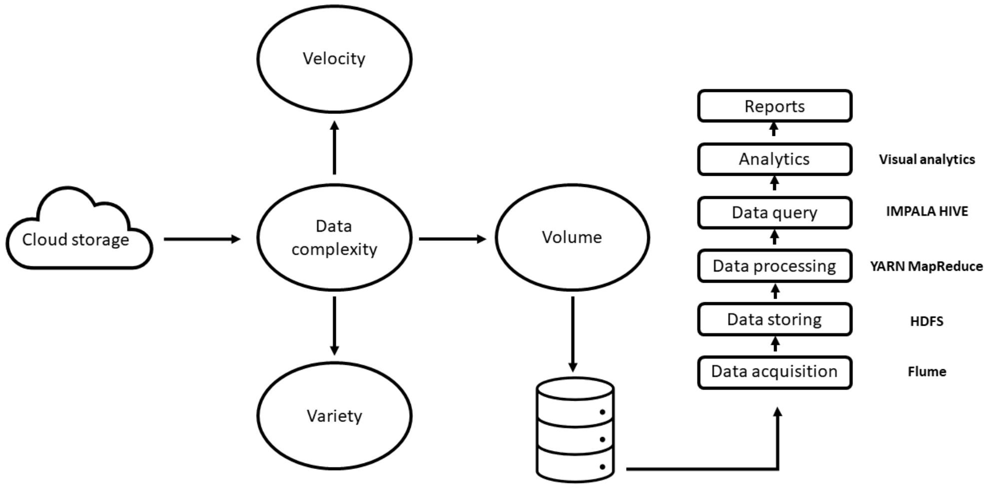



Figure 2.2: Big data and analytics.

example, cyber-physical systems continuously generate a large amount of data, which requires the big data techniques to process and help to improve system scalability, security, and efficiency to achieve the full autonomy in Industry 4.0, and the processed information is used for future business decisions.

The three Vs of big data, namely volume, velocity, and variety, as shown in Figure 2.2, are explained in the following.

**Volume:** Volume refers to the massive amount of data generated, which causes datasets to become too large for standard database technology. Larger data units, such as terabytes, petabytes, and exabytes, are used to measure this type of data [28].

**Velocity:** The velocity of data is the rate at which it is generated, processed, and transferred in real time [28].

**Variety:** The type of data (nature of data) varies, indicating whether the data are structured or unstructured [28].

These three data characteristics determine whether the stream of data can be fairly called big data. There are steps to handle the big data and extract meaningful information from them. For the processing of big data, a higher computing power is also needed.

## 2.4 Local and cloud computing

Computing capability that is confined to a single space can be called local computing. The servers used in local computing are the units of computing hardware that are dedicated

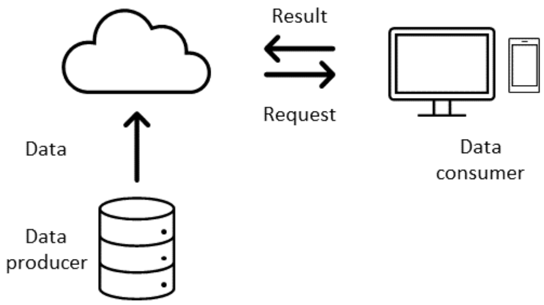



Figure 2.3: Cloud computing.

to performing specific tasks within a given physical space of a company. These servers are built based on the specific needs as indicated in [29]. For example, the data network used in a university could be considered local computing that meets its data processing requirements. The term cloud computing, in turn, refers to both the applications delivered as services over the Internet and the hardware and systems software in the large data centers that provide those services. Several computer hardware and software that offer some specific services to the general public result in what is commonly called cloud, as shown in Figure 2.3.

## 2.5 Edge computing

The integration of different digital technologies into industrial processes describes the IR4.0, resulting in increased automation, connectivity, and data-driven decision-making. Edge computing is critical in enabling and improving these capabilities. Edge computing is essential for the IR4.0, because it enables real-time decision-making and customization while ensuring a low latency, reliable operation, effective bandwidth utilization, data privacy, and scalability. Industries may fully benefit from the IR4.0 and increase their efficiency, productivity, and competitiveness by incorporating edge computing into their infrastructure [28, 30].

As its name entails, edge computing is the data processing paradigm that brings the data processing closer to the edge of the data network, i.e., closer to the end-application. In other words, the aim of the edge computing is that computation should be done in the proximity of data sources [31]. Figure 2.4 depicts the process of edge computing, where nodes at the network edge are performing several different tasks like computation, caching, device management, and even training of machine learning algorithms. This approach has the benefit of reducing the data traffic that will go through the Internet to reach the cloud.

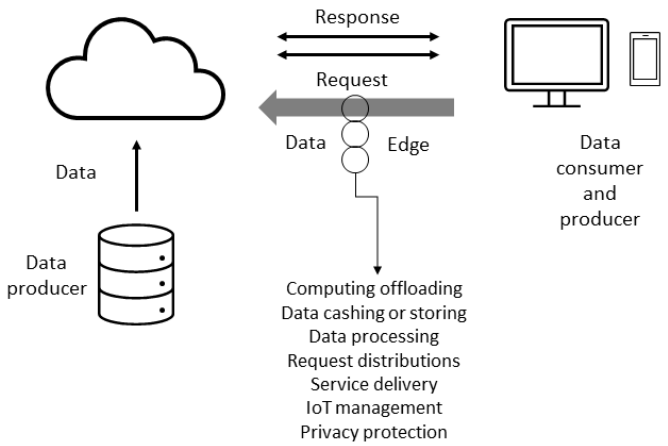



Figure 2.4: Edge computing, adapted from [32].

It is also emphasized that the decision between cloud and edge is nuanced. Most of the industries have shifted their operations to some form of cloud computing together with the increasing number of IoT devices with the huge amount of data acquired [33]. In the industrial environment, the increase in this vast amount of data from devices creates many problems. First, the huge amount of data has to be analyzed, and much of those data may be irrelevant to the operations. This creates high traffic to the central repository and a high cost of extra unnecessary storage. Second, the huge amount of data takes more time to send to the cloud server. This is crucial in many industrial applications, because for some applications, the difference between seconds and milliseconds can be very important. Third, sending data to the cloud and retrieving those data can be very costly [33].

Edge computing filters out the important data and analyzes them locally in real time, which may improve the speed of data analysis and also the decision-making process [33].

Table 2.1: Differences between the cloud and edge computing paradigms (adapted from Publication VI).

| Point of difference | Cloud computing | Edge computing |
|---|---|---|
| **Operations** | Happens on the cloud platforms, such as AWS, Azure, Google | Happens on the device itself |
| **Benefits** | Can store a massive amount of data on scalable hosting in the cloud, which can be accessible anytime on the Internet | Network can be scalable independently with each new device that is added to the system |
| **Suitable use cases** | Suitable for the operations with extreme latency concerns | Suitable for the businesses that require a huge amount of data storage and need scalable and cost-effective hosting providers |

In [34], a proof-of-concept platform was built for running a facial recognition application; the response time in the cloud was recorded to be 900 ms, and by shifting to the edge the response, the time reduced to 169 ms. In another study [35], cloudlets were used to offload computing tasks for wearable cognitive assistance, there were significant improvements in the response time, from 80 to 200 ms, and the energy consumption was reduced by 30%–40% by cloudlet offloading.

Although edge computing provides some benefits, this does not mean that the importance of cloud computing can, or should, be ignored, because having a centralized location for the data storage, process, and analysis still has many benefits. Edge computing can be used to process sensible information or time-sensitive data, allowing other less critical data to be processed in the cloud. For example, some data in the cloud may not be needed immediately but can be later used for machine learning purposes for the improvement of industrial operations and strategies [36]. A summary of the differences is presented in Table 2.1, which was first reported in Publication VI.

## 2.6 Cyber-physical system

The term cyber-physical system (CPS) emerged as one element of the IR4.0. CPS is associated with the automation and integration of industrial processes, which enable innovative functionalities using new capabilities related to data exchange and processing over networks. According to [37], CPSs comprise interacting digital, analog, physical, and human components engineered to function through integrated physics and logic. CPSs are mostly involved in the digitalization processes of various sectors of life.

In a general way, a CPS consists of a physical process, sensing devices (sensors/actuators), communication technology, data processing capabilities, and a decision-making mechanism. According to [7], the physical process can be digitalized by using sensing devices to get the data, and those data are then be sent to the data processing unit using the communication network. After the data processing, the information is shared either with humans or machines for decision-making as shown in Figure 2.5. The term CPS was coined by Helen Gill in 2006, and some key events about the CPS were organized at the National Science Foundation (NSF) workshop in October 2006 and later in the same year in the workshop on Network Embedded Control for CPS [3]. In the literature, with the passage of time, the definition of CPS has been based on six key characteristics, namely *hybrid system, hybrid methods, control, component classes, time,* and *trustworthiness* as shown in Table 2.2.

CPSs necessarily combine both computational and physical capabilities that can also involve humans, leading to new types of interactions between the physical world through computation, communication, and control.


Figure 2.5: Cyber-physical system, adapted from [7].

Table 2.2: Different definitions of CPS as presented in [3].

| Source | Characteristics | Definition |
|---|---|---|
| [38] | Hybrid Systems | “Cyber-Physical Systems (CPS) are integrations of computation with physical processes.” |
| [39] | Hybrid Systems | “CPS addresses the close interactions and feedback loop between the cyber components, such as sensing systems, and the physical components.” |
| [38] | Hybrid Methods | “The intellectual heart of CPS is in studying the joint dynamics of physical processes, software, and networks.” |
| [40] | Hybrid Methods | “The dynamics among computers, networking, and physical systems interact in ways that require fundamentally new design technologies.” |
| [41] | Control | “Cyber-Physical Systems (CPS) integrate computing and communication capabilities with monitoring and control of entities in the physical world.” |
| [38] | Control | “Cyber-physical systems (CPS) are physical and engineered systems whose operations are monitored, coordinated, controlled, and integrated by a computing and communication core.” |
| [42] | Component Classes | “CPS is envisioned to be a heterogeneous system of systems, which consists of computing devices and embedded systems including distributed sensors and actuators. These components are inter-connected together.” |
| [3] | Component Classes | “The computational and physical components of such systems are tightly interconnected and coordinated to work effectively together, sometimes with humans in the loop.” |
| [42] | Time | “These concerns are of particular importance in cyber physical systems in which computation and communication timing and event semantics are independent with physical timing and event semantics.” |
| [3] | Trustworthiness | “Cyber Physical Systems (CPS) are smart networked systems with embedded sensors, processors and actuators that are designed to sense and interact with the physical world (including the human users), and support real-time, guaranteed performance in safety-critical applications. ” |
| [43] | Trustworthiness | “A cyber-physical system (CPS) integrates computing, communication and storage capabilities with monitoring and/or control of entities in the physical world, and must do so dependably, safety, securely, efficiently and real time.” |

# 3 Publication overview

In this chapter, the publications included in this doctoral dissertation are presented. The chapter provides a summary of the research work carried out to achieve the main objective of the dissertation. For each publication, the motivation, the main context, the link to the research objectives, the general and specific research questions, the method used, and the findings and contributions are presented. This dissertation contains six research articles (Publications I–VI) published in different journals and conferences. The main topics covered in these publications are the role of the IoT and big data in developing cyber-physical systems, designing a theoretical framework for the selection of IoT platforms for industrial applications, highlighting the key factors of IoT platforms, and benefits of using the IoT in different industrial applications. A schematic presentation of the relation between Publications I–VI and the three research questions introduced in Section 1.3 is given in Table 3.1.

Table 3.1: Publications I–VI, their titles, and the associated research questions.

| **Publication** | **Title** | **Related RQ** |
|---|---|---|
| Publication I | Twenty-One Key Factors to choose an IoT platform: Theoretical Framework and its Applications | RQ1 |
| Publication II | Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data | RQ3 |
| Publication III | Operation of Power-to-X-Related Processes Based on Advanced Data-Driven Methods: A Comprehensive Review | RQ3 |
| Publication IV | IoT framework and requirement for intelligent industrial pyrolysis process to recycle CFRP composite wastes: application study | RQ3 |
| Publication V | Unified Framework to Select an IoT Platform for Industrial Energy Management Systems | RQ2 |
| Publication VI | Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing | RQ3 |

## 3.1 Publication I

**Title of Publication I:** Twenty-One Key Factors to choose an IoT platform: Theoretical Framework and its Applications.

### 3.1.1 Research objectives and the research question answered

The current trend of digitalization is bringing a huge change in the industrial sector, different physical elements, and processes connected to share information. Many industries have already started to shift part of their operations to the cloud. The IoT has played a major role in this, emphasizing the necessity of the proper selection of a suitable IoT platform based on the specific characteristics of the applications to be digitalized.

Currently, there are plenty of IoT platforms on the market, most of them having similar functionalities, but with different implementations and underlying technologies. Hence, selecting a suitable IoT platform for a given company is a hard task. There are multiple reasons behind this. One reason is the lack of experience in the management of the IoT and its platforms. A consequence of this is the lack of understanding of the company's own requirements in terms of technical specifications for the future IoT deployment.

The specific aim of Publication I was to first highlight the key building blocks of the IoT for the understanding of its functionality and significance by identifying and verifying the key factors of an IoT platform. The publication addresses RQ1 of the dissertation by answering the following specific research question:

***How to select a suitable IoT platform for industrial applications?***

### 3.1.2 Rationale and context

Publication I presents a concise overview of the need for understanding the IoT and its various components, as well as IoT platforms and the key factors associated with them. To grasp the functionality and significance of the IoT, it is crucial to comprehend its fundamental elements, which collaborate to deliver its functionalities and accomplish specific objectives. This study acknowledges six primary building blocks of the IoT, namely Identification, Sensing, Communication, Computation, Services, and Semantics.

IoT platforms play a vital role in providing essential services and features to IoT applications. An IoT platform constitutes a significant part of an IoT solution. When implementing the IoT for business applications and other functionalities, a company needs to select an appropriate IoT platform. There is a hike in the industrial digitalization, and companies are shifting their businesses to the cloud. However, there is a gap in the understanding of the IoT and IoT platforms. There is a vast array of IoT platforms available on the market, and finding one that adequately meets the company's business requirements poses a challenge. This predicament can be resolved by understanding the key factors of IoT platforms and aligning them with the company's business needs.

In this study, 21 key factors of IoT platforms were identified from the literature. Additionally, the importance of these factors was validated by a survey to IoT experts. Moreover, based on the discussion, a theoretical framework was proposed for the selection of the IoT platform for companies.

### 3.1.3 Research method and data collection

In the study, 21 key factors of an IoT platform were collected, and the importance of those services was verified by a questionnaire survey sent to experts in different universities. The data collected for this study are from research articles published in different scientific journals and conferences. Various websites were studied, including the websites of different IoT platforms. Initially, a large number of research articles and websites were studied, multiple articles addressing the same topics were removed, and finally, the sample was limited to 46 studies.

In the research, a two-round Delphi study was conducted. Based on the literature review, a questionnaire was designed consisting of 21 questions, each of which was related to the IoT and the IoT platform as shown in Table 3.2. For the answers, a five-point scale from 1 to 5 was used (1 = "totally agree", 2 = "agree", 3 = "neutral", 4 = "not agree", and 5 = "totally disagree"). Fifteen IoT experts from three different universities were selected for the survey, the response time was two weeks, and the questionnaire was sent by email.
In the first round, 14 experts replied to the questionnaire as shown in Figure 3.1. The response rate of the first round was relatively high, and the experts (80%) mostly agreed with the questions. A few experts (6%) disagreed, and some of the experts (14%) gave a neutral response to some questions.

During the second round of the survey, the same questionnaire was used, however, with a summary of the experts' responses of the first round and some explanations of the questions included. Again, the questionnaire was sent to the 15 experts by email, and the re-

Table 3.2: Questions used in the survey, applying the Delphi method (Publication I).

| Q# | Survey question |
|---|---|
| Q1 | What is your opinion about the importance of stability of IoT platform? |
| Q2 | What is your opinion about the importance of Scalability of the enterprise of IoT platform? |
| Q3 | Do you think that IoT platform should be flexible with the advancement of technologies? |
| Q4 | Do you think it is important to know about the pricing models before selecting IoT platform? |
| Q5 | Do you think IoT platform should provide security at both the ends, software and hardware? |
| Q6 | Do you think IoT platform can reduce Time to market for the business? |
| Q7 | Do you think IoT platform should support the basic descriptive, predictive and perspective analytics? |
| Q8 | Do you think it is important to know who will own the data collected by IoT platform? |
| Q9 | Is it important to know the application environment of IoT platform? |
| Q10 | Do you think it is important to know the Ownership of cloud infrastructure? |
| Q11 | Do you think extend of legacy architecture in IoT platform is important? |
| Q12 | Do you think Edge intelligence is important for IoT platform? |
| Q13 | Do you think IoT platform needs high bandwidth networking? |
| Q14 | Do you think it is important for IoT platform to support new Protocols and its updated versions? |
| Q15 | Do you think the IoT platform vendors should implement some steps to keep System performance high? |
| Q16 | Do you think the providers should have some dedicated infra to handle customer data if there is some problem in IT? |
| Q17 | Do you think Hybrid cloud is important for IoT platforms? |
| Q18 | Do you think the providers should offer facilities to customers for any possible migration to other IoT platform? |
| Q19 | Do you think IoT platform Interoperability will enable the organization to get higher productivity? |
| Q20 | Is it necessary to check the previous experience of IoT platform, before selection? |
| Q21 | Is it necessary that user interface of the IoT Platform should be simple and attractive? |

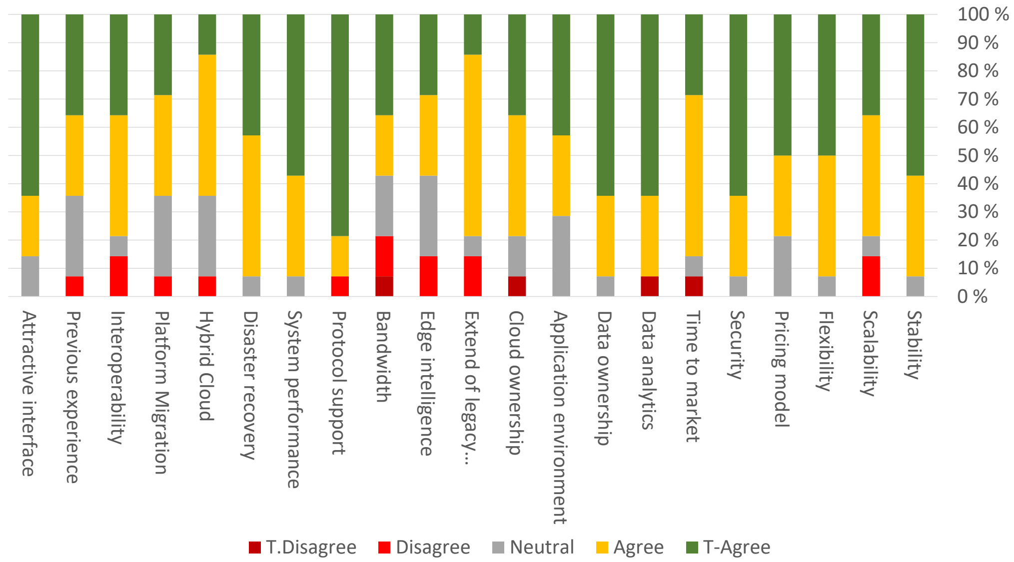



Figure 3.1: Expert opinions in the first round (Publication I).

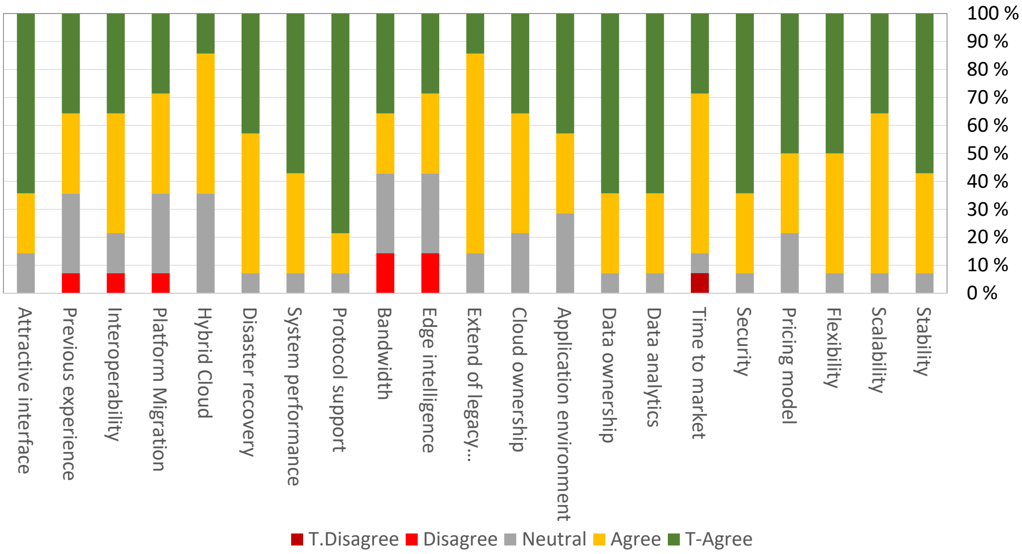



Figure 3.2: Expert opinions in the second round (Publication I).

sponse time was two weeks. In the second round, 14 experts replied to the questionnaire. Some of the experts had changed their opinion about some questions; the percentage of “agree” changed from 80% to 81%, the percentage of “disagree” changed from 6% to 4%, and the percentage of “neutral” increased from 14% to 15%. The results are shown in Figure 3.2.

Table 3.3: Results of both the rounds of the Delphi study. The mean and median are taken from the “agree” values (Publication I).

| | | **Survey round 1** | | | | | **Survey round 2** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **F** | **Factor** | Mean | Median | Disagree % | Neutral % | Agree % | Mean | Median | Disagree % | Neutral % | Agree % |
| F1 | Scalability | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 0% | 7% | 93% |
| F2 | Flexibility | 4 | 4.5 | 0% | 7% | 93% | 4 | 4.5 | 0% | 7% | 93% |
| F3 | Data analytics | 4 | 5 | 7% | 0% | 93% | 5 | 5 | 0% | 7% | 93% |
| F4 | Disaster recovery | 4 | 4 | 0% | 7% | 93% | 4 | 4 | 0% | 7% | 93% |
| F5 | Stability | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F6 | Security | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F7 | Data ownership | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F8 | Protocol support | 5 | 5 | 7% | 0% | 93% | 5 | 5 | 0% | 7% | 93% |
| F9 | System performance | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F10 | Time to market | 4 | 4 | 7% | 7% | 86% | 4 | 4 | 7% | 7% | 86% |
| F11 | Legacy architecture | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 0% | 14% | 86% |
| F12 | Attractive interface | 5 | 5 | 0% | 14% | 86% | 5 | 5 | 0% | 14% | 86% |
| F13 | Pricing model | 4 | 4.5 | 0% | 21% | 79% | 4 | 4.5 | 0% | 21% | 79% |
| F14 | Cloud ownership | 4 | 4 | 7% | 14% | 79% | 4 | 4 | 0% | 21% | 79% |
| F15 | Interoperability | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 7% | 14% | 79% |
| F16 | App. environment | 4 | 4 | 0% | 29% | 71% | 4 | 4 | 0% | 29% | 71% |
| F17 | Hybrid cloud | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 0% | 36% | 64% |
| F18 | Platform migration | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 7% | 29% | 64% |
| F19 | Previous experience | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 7% | 29% | 64% |
| F20 | Edge intelligence | 4 | 4 | 14% | 29% | 57% | 4 | 4 | 14% | 29% | 57% |
| F21 | Bandwidth | 4 | 4 | 21% | 21% | 57% | 4 | 4 | 14% | 29% | 57% |
| – | **Percentage** | – | – | **6%** | **14%** | **80%** | – | – | **4%** | **15%** | **81%** |

Table 3.3 shows the results of both rounds with some statistical calculations. For simplicity, “agree” and “totally agree” are merged into “agree”, and “disagree” and “totally disagree” into “disagree”.

### 3.1.4 Results and conclusions

The results of both the rounds of the survey are shown in Table 3.3. In the first round, the percentage of “agree” was 80% , the percentage of “disagree” was 6%, and the percentage of “neutral” was 14 %. In the second round, the “agree” percentage increased to 81%, the “disagree” percentage reduced to 4%, and the “neutral” percentage reached 15%.

The top five IoT platforms based on their market share are considered in the theoretical five-stage methodology of this study for choosing an IoT platform for industrial applications (see Figure 3.3). The company outlines its business requirements in the first round and explains how important IoT platforms and components are for the company. In the second round, the requirements are prioritized by the company as required (R), important (I), and not necessary (-). The third step compares the features offered by the IoT platforms to the business requirements. The fourth stage involves choosing IoT platforms that satisfy all of the company’s business criteria or the majority of them. There is a chance that several IoT systems are able to meet the requirements. In the fifth stage, based on a comparison of their match to I variables, including, e.g., pricing and time to market, a company can choose an IoT platform that is appropriate for its business goals. Additionally, it is possible that none of the platforms offer all the necessary functionalities. In that

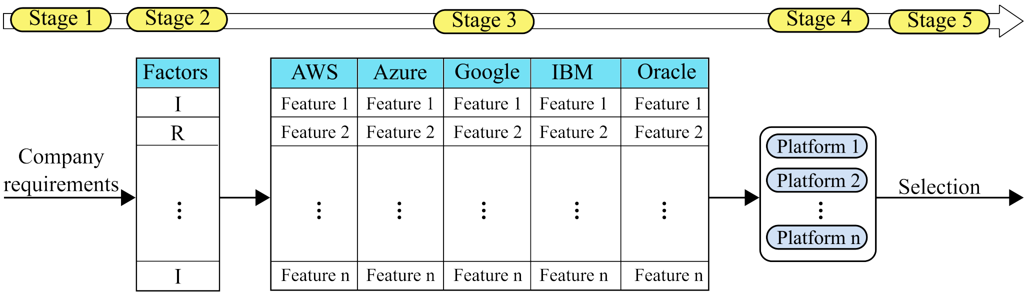



Figure 3.3: Theoretical five-stage framework for the IoT platform selection (Publication IV).

case, the company can decide to study more IoT platforms.

The purpose of Publication I was to develop an objective technique that will assist enterprises in choosing the best IoT platform based on their unique business requirements. To this end, the IoT's fundamental components were addressed, and it was explained how they work together to carry out particular activities. Second, based on the literature, 21 essential components of IoT platforms were identified, which were then used in the Delphi study to be confirmed with industry experts. The theoretical framework developed for choosing an IoT platform was tested using five well-known IoT platforms. Publication I offers thus a broad framework for choosing the best IoT platform for a particular firm by evaluating its unique requirements in comparison with the capabilities provided by the various platforms.

## 3.2 Publication II

**Title of Publication II : Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data.**

### 3.2.1 Research objectives and the research question answered

An industrial energy management system (IEnMS) can offer several approaches for changing the profile of the energy demand of companies using ICTs, sensors, and actuators. For instance, an industrial process can be programmed to run automatically when solar photovoltaic (PV) electricity is available. The IEnMS can also track the amount of energy used by various industrial processes and machines, allowing it to identify undesirable patterns and alert the appropriate staff to replace them with more energy-efficient models.

The main goals of this study were to identify the existing solutions, highlight their key requirements, and suggest a high-level architecture for the IEnMS that takes into account the IoT, big data processing, and data analytics while considering relevant aspects collected in a survey to experts employed in the relevant industry. Publication II then serves as an in-depth example that solves RQ3 by answering the following specific research question:

***How do the IoT, big data, and analytics support an energy efficient operation in the industry by the IEnMS?***

### 3.2.2 Rationale and context

Publication II provides the details of the IEnMS. Like in Publication I, the IoT and its components, IoT platforms, and key factors are addressed, and a framework for the selection of IoT platforms is designed. Publication II focuses on the IEnMS and its components, practices, and processes.

An IEnMS is desirable to ensure the successful implementation of the company's energy management. This allows the development of an objective method for monitoring, planning, and regulating the company's energy consumption and efficiency, and the achievement of the planned energy-positive goals.

### 3.2.3 Research method and data collection

In this study, a quantitative research method was used by collecting research data from previous publications in which the main components like *planning/strategy, operation/ implementation, control, organization, and culture* of the IEnMS have been identified. The data were later verified by a detailed questionnaire to experts in ten large companies to improve the performance of the IEnMS by using the latest technologies like the IoT, big data, and data analytics in the IEnMS.

The data collected in this study were based on a survey, and a questionnaire was designed that contains 20 questions with five options: "totally disagree", "disagree", "neutral", "agree", and "totally agree" as shown in Table 3.4 with some statistics about the expert opinions. Representatives of ten large companies replied to the survey.

### 3.2.4 Results and conclusions

The numbers of respondents choosing "totally disagree" or "disagree" were lower. Some of the company experts did not respond to some of the questions, such as those about the development of an energy management team by the energy manager and the lack of management awareness of the IEnMS. In general, the majority of respondents expressed that they "totally agreed" or "agreed" with the majority of the points, such as whether or not the industries should use the IEnMS to take advantage of the newest technologies like the IoT, big data, and data analytics, and whether or not this will help the industries in terms of reduced energy consumption, efficient energy use, lower energy bills and costs, and lower greenhouse gas emissions.

Table 3.4: Survey questions (Publication II).

| S.No | Topic | Total Disagree | Disagree | Neutral | Agree | Totally Agree | Mean | Median |
|---|---|---|---|---|---|---|---|---|
| Q1 | An Industrial Energy Management System is good for energy saving | 0% | 0% | 0% | 30% | 70% | 4.7 | 5 |
| Q2 | An Industrial Energy Management System can reduce greenhouse gas emissions | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q3 | There is lack of management awareness in industries for energy management | 0% | 20% | 30% | 30% | 20% | 3.5 | 3.5 |
| Q4 | An Energy Management System will provide energy saving opportunities | 0% | 0% | % | 50% | 50% | 4.5 | 4.5 |
| Q5 | The company's energy policy plays an important role in designing energy management system | 0% | 10% | 10% | 20% | 60% | 4.3 | 5 |
| Q6 | Long term energy planning is important for industries | 0% | 0% | 0% | 30% | 70% | 4.7 | 5 |
| Q7 | An energy manager should create an energy management team | 0% | 0% | 30% | 40% | 30% | 4 | 4 |
| Q8 | It is important for the energy manager to be environment friendly | 11% | 0% | 11% | 11% | 67% | 4.2 | 5 |
| Q9 | It is important that companies should display department-wise energy usage status using screens to motivate staff members to save energy | 0% | 0% | 20% | 40% | 40% | 4.2 | 4 |
| Q10 | It is important for companies to give incentives and rewards to staff to encourage them to achieve energy targets | 0% | 10% | 20% | 20% | 50% | 4.1 | 4.5 |
| Q11 | Companies need some renovation in the existing infrastructure to improve energy management infrastructure | 0% | 0% | 22% | 22% | 56% | 4.3 | 5 |
| Q12 | It is important for companies to add green energy like solar and wind energy to their existing energy usage | 0% | 0% | 10% | 20% | 70% | 4.6 | 5 |
| Q13 | Companies should have a strong policy to reduce greenhouse gas emissions | 0% | 0% | 10% | 0% | 90% | 4.8 | 5 |
| Q14 | Installing sensors on machines, so that these machines can use IoT-based techniques to share real data with each other, will lead to improved energy efficiency and performance | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q15 | Using IoT and Big data in Industrial Energy Management Systems will facilitate timely identification and prevention of faults | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q16 | Companies should invest more in Industrial Energy Management Systems (IEnMS) | 0% | 0% | 10% | 60% | 30% | 4.2 | 4 |
| Q17 | IoT is helping to improve HVAC (heating, ventilation, and air conditioning) systems in manufacturing plants. | 0% | 0% | 20% | 50% | 30% | 4.1 | 4 |
| Q18 | IoT devices are capable of collecting a huge amount of real time data about different machines. Therefore, collecting and using Big Data is a good option for companies to perform real time data about different machines. | 0% | 0% | 11% | 33% | 56% | 4.4 | 5 |
| Q19 | IoT and Big Data will make data analysis and processing easier and will give energy information very quickly and this can be useful for business decisions in the future. | 0% | 0% | 11% | 22% | 67% | 4.6 | 5 |
| Q20 | Industries should use the latest IoT-enabled technologies in their Energy Management System to improve energy activities like efficiency, performance, usage, cost etc. | 0% | 0% | 10% | 20% | 70% | 4.6 | 5 |

Figure 3.4 shows the results (with different colors) of the questions asked from the experts. Here, the purple color indicates the industry expert opinions "totally agree", light gray "agree", light orange "neutral", red "disagree", and blue "totally disagree". The majority of the companies recognized the importance of the points asked in the 20 questions. The companies agreed about the importance of the IEnMS in terms of energy efficiency, energy consumption, and reductions in energy costs and greenhouse emissions by using renewable energy and the latest technologies like the IoT, big data, and analytics.

Table 3.4 shows the survey questions and the response percentages of the experts. A design for the IEnMS was presented based on the most recent technologies, the IEnMS needs, and the opinions of the company experts as shown in Figure 3.5. The design consist of five stages. The first stage is data collecting, during which large amounts of data are gathered from various sources, including, e.g., machinery, renewable energy sources,


Figure 3.4: Survey results by topics, in percentages (Publication II).

and CCTV. This volume of data is kept in an inexpensive cloud storage. The second stage is data acquisition, during which the gathered data are stored in the master nodes of the Hadoop cluster, which can accommodate numerous data types produced by heterogeneous devices. The incomplete data are either fixed or eliminated throughout the data preparation process. The data are then saved in Flume in one or more channels. In the third stage, the stored data are sent to the Hadoop Distributed File System (HDFS) repository, where they are formatted as needed. The large file is divided into several blocks by the HDFS, and these blocks are stored in various data nodes. The data that are stored in the HDFS are examined by Yet Another Resource Negotiator (YARN). In the fourth stage, HIVE is used to run SQL queries on the stored data in the HDFS. The fifth stage is the last stage, in which the calculated data (information) are shared with the energy management for, e.g., future planning and decision-making.

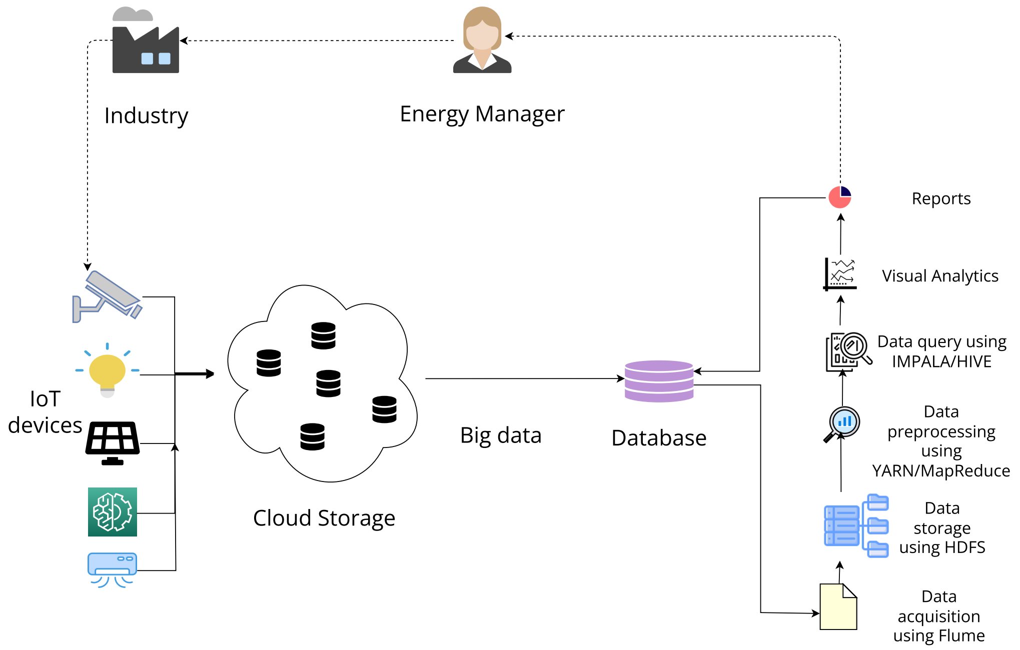



Figure 3.5: Data generation using the IoT and processing with big data. The solid arrows indicate the data flow from one element to another, and the dashed line shows possible information about the industrial operations based on the data reports from the IEnMS (Publication II).

Publication II emphasizes the value of the IEnMS and describes each of its parts in detail. The IEnMS outlines the procedures and techniques used in industrial energy management, which are often regarded as supporting activities. The study discussed the advantages that the IoT, big data, and data analytics provide to the industrial sector as well as why the IEnMS should use them. In the study, large companies were surveyed in detail, and the findings indicate that the majority of industrial professionals support the use of contemporary technologies in the IEnMS. A theoretical framework was developed for acquiring energy information by utilizing the IEnMS and contemporary technologies based on the advice provided by industry experts. In this framework, IoT devices are used to gather machine data, which are then sent to a database, where the big data process and data analytics are started. Finally, the data-generated information is transmitted to the energy manager, a specialist in energy management. By using this strategy, businesses should be able to reduce greenhouse gas emissions while also improving their energy efficiency. Additionally, the information provided by the data can be used for industrial equipment maintenance as well as for present and upcoming business decisions.

A survey administered to energy experts in various companies (all the ten companies surveyed) found that improvements in energy efficiency are anticipated to be made by using the IEnMS as a centralized entity that will support the business in making better strategic and operational decisions based on data, while enabling an energy-centric operation through specialized IoT devices, sensors, and actuators.

## 3.3 Publications III–VI: Cases of CPS and benefits of the IoT

This section highlights the benefits of the IoT in different industrial applications to be operated as a CPS. It also explains how to use the proposed IoT platform selection methodology (proposed in Publications I, II) in different industrial applications. The section is built based on the research work of Publications III–VI to answer RQ2 and RQ3. Publication V is used to answer RQ2. The specific research questions of these papers are presented case by case.

### 3.3.1 Publication III: Operation of Power-to-X cogeneration plants based on advanced data-driven methods: A comprehensive review.

The main objective of this study was to propose a data-driven approach that enables the cyber-physical operation of power-to-X industrial plants. This study is based on the work done in Publication II, in which the key building blocks of the IoT are discussed, and the operation and principles of the IoT and big data are explained. In this context, the specific research question is given as follows:

***How to use the IoT in power-to-X-related processes based on advanced data-driven methods?***

When introduced, the idea of producing renewable, carbon-neutral fuels that absorb $CO_2$ during production sounded unrealistic. Now, however, this technology is a fact. One of the primary strategies to accelerate the energy transition will be P2X, especially when coupled with sector coupling. P2X collects $CO_2$ from the atmosphere and combines it with green hydrogen to produce a variety of future fuels that are carbon-neutral. Publication III highlights the importance of data-driven methods used for the collection of data from the various parts of the methanation reactor (e.g., electricity source, electrolysis, $CO_2$ capture, hydrogen $H_2$ storage). In this study, the IoT was used for data collection from the overall process of a methanation reactor. The collected data are stored in an inexpensive cloud storage, from where the data are transferred to the big data for data analysis, calculations, and provision of information. Later on, the same calculated data are passed to the machine learning process for the future predictions, and based on them, to the future planning and decision-making process. Based on the facts and the current high-level technologies, a theoretical framework was designed, which will help in all the processes starting from data collection to the future predictions by using machine learning.

Publication III employed a quantitative research method, and the data used in the research

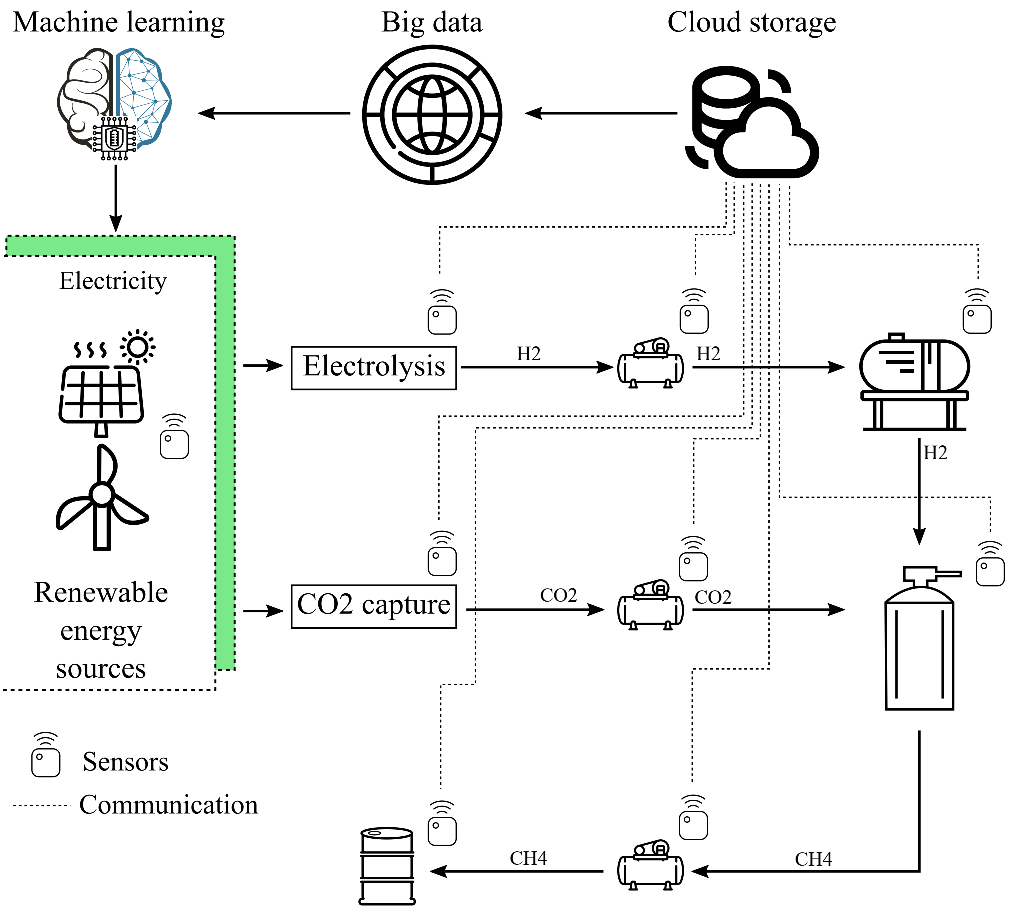



Figure 3.6: Operation of the methanol synthesis. The proposed architecture (Publication III).

were produced with a Matlab model by one of the coauthors of this publication. Based on the literature review and the requirements of methanol synthesis, an architecture was designed for the data gathering from the whole process, and later, the data were stored in a cloud storage. From there, the data are sent to big data and analytics, from where the data are available for the machine learning process.

In Publication III, an advanced architecture was designed for a P2G system including advanced data-driven methods, such as the IoT, big data, and machine learning as shown in Figure 3.6. In this study, an IoT-based system (architecture) was designed that consists of four interconnected stages.
The IoT sensor devices are integrated into the entire physical infrastructure (renewable energy sources, the electrolysis process, and hydrogen and methane storage) of the P2X process at the first stage in order to collect data. The information collected at this stage comes from various components of the proposed architecture, such as solar panels, wind

turbines, air pressure, atmospheric temperature, electricity price, the amount of electricity used for the electrolysis process and $CO_2$ capture, the amount of $H_2$ produced and stored, the amount of $H_2$ and $CO_2$ used to produce methanol, and the amount of methanol itself. The data generated in the first stage are transmitted and saved in a cloud storage in the second stage, where they are processed further. This stage requires wired or wireless internet connectivity, or in some situations, access to nearby computer units that use edge computing or private networks.

In the third stage, the large amount of data that has been stored is sent to big data analytics tools for additional processing. All types of data, including structured, semistructured, and unstructured data, are included in the big data. Opportunities to enhance industrial services are then created by the data that the IoT services generate. For enhanced information presentation as part of an expert system, the data produced by IoT sensors, for instance, can be examined in real time. This will help future decision-making and system operation. This third stage can be further divided into four phases, to be presented next, that make up the big data process.

The initial stage comprises the following tasks: Taking into account the requirements of the intended end application, massive data collected in the second stage are stored. The collected data are subsequently transmitted to the master node of the Hadoop cluster. Because of the diversity of data formats from heterogeneous devices, data preparation is required. Accurate and incomplete data are handled during data preparation, and incomplete data are either rectified or eliminated. Flume is used to collect data; it gathers, aggregates, and sends large amounts of data to the Hadoop master node. Flume records the information it receives through one or more channels.

In the second stage, the data are transmitted to an external Hadoop Distributed File System (HDFS) repository. Then, by storing individual blocks of big files in a number of data nodes connected to the master node, data are serialized and written in a specific format. Any sort of data, whether structured, unstructured, or semistructured, can be stored using the HDFS. DataNodes make up the HDFS clusters. These DataNodes store both the file system metadata and the actual data together. It is possible to handle jobs on nodes where the data are present because the two run on the same set of nodes. Data saved in the HDFS are analyzed using YARN.

In the third stage, by using the tools Hive and Impala, SQL queries can be run on the HDFS data. HIVE is specifically used for data selection, analysis, and computation on the relevant data.

Data analytics is the last step, releasing the data that have been processed for use as a decision-support tool. Hadoop data analytics is carried out by using Scalable Advanced Massive Online Analysis (SAMOA), a distributed streaming machine learning system. Finally, the prediction stage is the fourth stage. This stage involves giving the machine learning or expert system the cleaned-up data from the big data repository, either as a training set or a decision-support tool.

Publication III examined the application of current advances in ICT, including the IoT, big data, and machine learning, in P2X processes. The study specifically focused on the production of renewable electricity that would profit from the utilization of the most recent P2X technology. A further objective was to study how to efficiently deploy recent ICTs to support the operation of P2X applications. In this regard, the paper discussed the collection, distribution, and processing of data in P2X scenarios as well as the available technology for putting such CPSs into practice.

### 3.3.2 Publication IV: IoT framework and requirement for intelligent industrial pyrolysis process to recycle CFRP composite wastes: application study.

This contribution focuses on the digitalization of a pyrolysis process in industry. The study is based on the work done in Publication I, in which the key factors of an IoT platform are discussed, and a five-stage theoretical framework is designed to be used to select a suitable IoT platform for industrial application based on business requirements. This study answers the following research question:

***How to select an IoT platform for an intelligent industrial pyrolysis process?***

Publication IV addresses carbon fiber recycling. For decades, the utilization of carbon fiber-reinforced polymer (CFRP) composites in high-performance applications has increased tremendously. The composites can replace traditional metals in lightweight applications because of their outstanding mechanical qualities for a low weight ratio. However, 20 years later, CFRP composites still in use have reached the end of their useful lives (EoL), raising serious concerns about how to dispose of them. As of now, 62,000 tonnes of CFRP composite trash have accumulated annually, and a forecast states that if properly disposed of, the amount might rise to 90,000+ tonnes/year. The yearly demand for virgin CFRP composites is also anticipated to rise from 72,000 to 140,000 tonnes/year in the same period. The only sustainable solution to achieve a balance is to recycle the used composites, recover the precious carbon fibers (CFs), and incorporate them into new composites.

The pyrolysis recycling process to recover CFRP composite wastes is depicted in Figure 3.7. The composite waste is first mechanically shred to minimize its size before being fed into the system. The pyrolysis reactor is a sealed chamber without any oxygen. To separate the valuable CFs from the matrix, the procedure is carried out at 550 °C for the necessary amount of time (dependent on the amount of waste), but in an inert gas atmosphere. The resin leftovers are then oxidized by passing the recovered fibers through a secondary heating chamber at 200 °C. Finally, the CFs are recycled to produce hot gas and pyrolytic oil, which can both be utilized as feedstock. Overall, pyrolysis on an industrial scale offers huge long-term advantages.

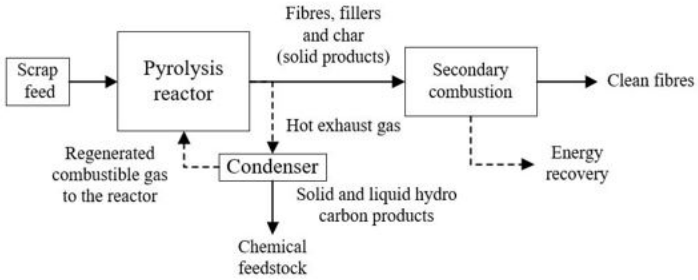



Figure 3.7: Pyrolysis process (Publication IV).

It is important to efficiently dispose of the accumulated carbon fiber-reinforced polymer (CFRP) composite trash. In wealthy nations like the UK and Germany, pyrolysis, the most efficient thermal-based recycling method to date, has progressed significantly toward industrial scalability. Within such a large operational environment (e.g., higher than 1 tonnes/day operating capacity), even the smallest errors can result in undesirable outcomes and delays in the workflow. To reflect the various classifications and volumes of the CFRP composite wastes, the existing semiautomated and, in some cases, fully automated plants should be regularly updated. An IoT-based approach was suggested to address these knowledge gaps and errors.

Using the concepts of cyber-physical systems, Publication IV investigated the theoretical integration of an IoT-based framework into the pyrolysis process to recycle CFRP composite waste. The proposed framework is made up of sensors and actuators that will be used to gather data and communicate with a central management built as a platform that will articulate and manipulate data to meet the needs of the recycling process, computationally modeled through logical relations between physical entities. The goal of integrating IoT technology into industrial-scale pyrolysis is to enhance its performance and factors like cost, speed, reliability, and scalability. Additionally, the publication addressed debates on the data acquired from the Internet of Things (IoT) as a trustworthy source of knowledge in the present decision-making and future advancements in the pyrolysis process.

Figure 3.8 shows the IoT network that was built for an industrial-scale pyrolysis setup. At every step that is critical to the process, from the waste feed to the CF recovery, sensors and actuators are used. The IoT framework is intended to be adaptable and expanded to a variety of furnaces (pyrolysis process 1, pyrolysis process 2, and pyrolysis process n). The setup for the pyrolysis process is shown in Figure 3.7. The primary gateway sends the data collected from each furnace to the cloud for storage, processing, and display. Additionally, information can be obtained through cloud-based data. Based on that information, a further business choice can be made in the future that will accurately and precisely improve and speed up the CF recycling process.

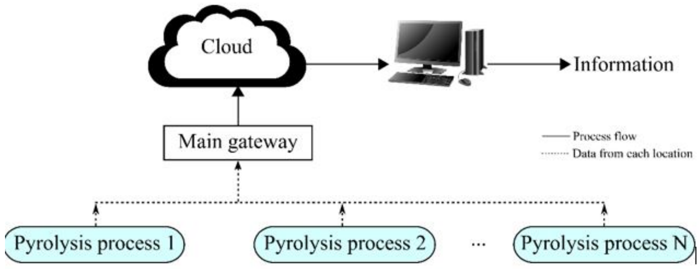



Figure 3.8: Implementation of the IoT framework in industrial-scale pyrolysis (Publication IV).

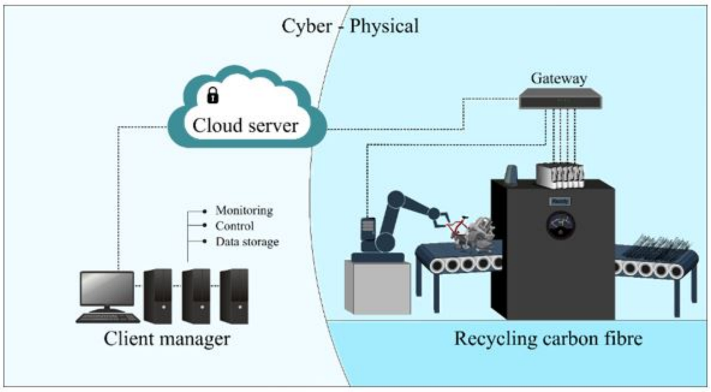



Figure 3.9: IoT schematic of an in-site CF recycling process (Publication IV).

By incorporating physical elements into the digital realm of what is referred to as a cyber-physical system (CPS), the CFRP waste recycling process will benefit from the deployment of the IoT. This will make it possible to better regulate the input material moved by the conveyor, control the furnace temperature, and monitor the creation of char in real time. As demonstrated in Figure 3.9, the IoT is being utilized to create a cyber-physical environment where actuators and sensors are connected to a gateway that serves as an interface to a cloud server. A manager can monitor, manage, or analyze stored data remotely in the digital environment.

At the end of the study, the framework for IoT platform selection presented in Publication I was used to select a suitable IoT platform for the pyrolysis process based on the requirements of the process. The prerequisites for the pyrolysis process aiming at recycling CFRP composite waste encompass the elimination of manual labor, acceleration of recycling rates (more than 1 ton per day), enhancement of energy efficiency, and improvement in overall efficiency. Additionally, the framework for IoT platform selection involves determining the necessary heat for processing a specific quantity of CFRP composite waste and estimating the heat generated during the process. Moreover, in the IoT platform selection process, emissions (both exhaust and external to the system) are calculated. In summary, the pyrolysis process necessitates stability, flexibility, scalability, security, an appealing user interface, data analysis capabilities, and system-wide interoperability. Furthermore, the process requires a user-friendly application development environment for its IoT cloud-based business application. These pyrolysis process requirements are divided into two categories: ”important” and ”required”. Among them, stability, flexibility, scalability, security, user-friendly interface, and data analysis are considered ”important”, while interoperability is deemed ”required”. The pyrolysis process requirements are then compared with the features offered by five major IoT platforms, as outlined in Table 3.5. All of the critical pyrolysis process requirements are aligned with the features offered by the AWS IoT platform. However, Azure falls short in terms of the necessary flexibility and the required interoperability factor.

Table 3.5: Features provided by IoT platforms [44]

| **Factors** | **AWS** | **Azure** | **Google** | **IBM** | **Oracle** |
|---|---|---|---|---|---|
| Scalability | Yes | Yes | Yes | Yes | Yes |
| Flexibility | Yes | - | Yes | - | Yes |
| Stability | Yes | Yes | Yes | - | - |
| Security | High | High | High | High | High |
| Data analytics | Yes | Yes | Yes | Yes | Yes |
| Disaster recovery | Yes | Yes | No | No | No |
| Data ownership | - | Yes | - | - | - |
| Protocol support | Yes | Yes | - | Yes | Yes |
| System performance | Yes | - | Yes | Yes | - |
| Interoperability | Yes | - | - | - | Yes |
| App. environment | Yes | Yes | Yes | Yes | Yes |
| Cloud ownership | Yes | Yes | Yes | | Yes |
| Pricing model | Bad | Bad | Good | - | - |
| Legacy architecture | Yes | - | - | - | Yes |
| Attractive interface | Yes | Yes | - | No | - |
| Time to market | Yes | Yes | | | Yes |
| Bandwidth | - | - | Good | - | - |
| Edge intelligence | Yes | Yes | Yes | - | Yes |
| Hybrid cloud | Yes | Yes | - | - | - |
| Platform migration | Yes | Yes | - | - | - |
| Previous experience | Yes | Yes | - | - | - |

### 3.3.3 Publication V: Unified Framework to Select an IoT Platform for Industrial Energy Management Systems

This publication is based on the work done in Publication I, in which the six main building blocks of the IoT are explained, and 21 key factors of IoT platforms are identified.

Moreover, Publication V also uses the key components and requirements of the IEnMS presented in Publication II. The goal is to answer RQ2. The main research question of the publication is the following:

***How to select an IoT platform for industrial energy management systems based on business requirements?***

By combining the IoT and big data, Publication V aimed to use the proposed theoretical framework to choose an appropriate IoT platform for the IEnMS. In order to help different sectors choose the best IoT platform for the IEnMS based on their unique demands and business requirements, an objective yet universal methodology was developed. In order to find the ideal match for their IEnMS, enterprises can use Publication V to help them conduct a thorough study of their own energy needs and understand the major elements of their energy management (EnM).

A computer-based system known as an industrial energy management system is used to gather and measure energy data from the user's point of use, such as HVAC (heating, ventilation, and air conditioning) units, lighting systems, and water and gas meters installed on production lines. Below, a more detailed description is given of how the IEnMS is working by using the IoT and big data.

**Build a data collection strategy:** A system that is used to gather precise real-time data at the granular level, as well as data on the location, timing, and type of energy consumption by various machines. On the incoming supply and the major energy consumer (device), sensors, smart meters, and submeters are installed in order to collect the data. Possible targets include, for instance, boilers, production lines, and HVAC systems. The objective of this section focuses on the data gathering in real time to identify the locations where most energy is consumed.

**Transform raw energy data into useful information:** In this step, the acquired data are processed, evaluated, and transformed into meaningful information. Here, big data technology is used to import raw data from various machines through IoT devices and then transform the raw data into valuable information in the form of user-friendly charts, graphs, and other visuals. Here, the raw data gathered can be tied to production levels, weather information, human behavior, and other variables that may have an impact on how much energy is used to produce the key performance indicators.

**Assign responsibility, analyze data:** The information provided in this stage must be transformed into reports that are applicable and insightful. This can be achieved by combining the information with knowledge about the facility and some energy management skills; for instance, the energy manager of the company can carry this out. The information of the industrial energy management system is interpreted by the energy manager, who then combines it with the business procedures to develop goals.

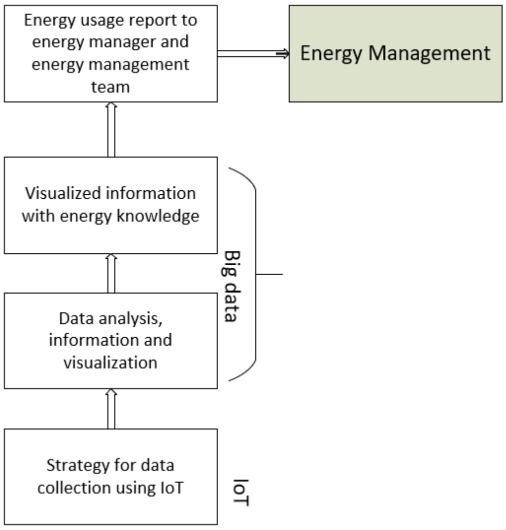



Figure 3.10: Operation of the IEnMS based on the IoT and big data (Publication V).

**Interpret the results and agree an action plan:** At this stage, the energy manager has access to the energy usage reports. To initiate an energy strategy and come to an agreement on an action plan, the energy manager and the energy management team get in touch with the departments.

After operation of the IEnMS based on the IoT and big data as shown in Figure 3.10, some requirements for the operation of the IEnMS were collected as shown in Table 3.6. The selected requirements were then compared with the features provided by the IoT platforms as discussed in Publication I to select a suitable IoT platform for the IEnMS business application.

Table 3.6: Requirements of Industrial Energy Management (Publication V).

| S No | Requirements for IEnM |
|---|---|
| 1 | Development and implementation of a plan that includes the energy policy and targets for savings. |
| 2 | Organizing different energy activities including the allocation of responsibilities and tasks. |
| 3 | Establishment of a team led by an energy manager, who will be responsible to report directly to the high management. |
| 4 | Development of policies and procedures, which can include, e.g., energy procurement, usage, and cheap purchases |
| 5 | Carrying out the initial energy audit to identify energy saving potentials |
| 6 | Planning and implementation of energy efficiency measures |
| 7 | Identification of the company-specific key performance indicators that can measure the progress on a regular basis. |
| 8 | Implementation of meters for monitoring of the energy consumption in the production processes at regular intervals. |
| 9 | Reporting of the information gathered from the data to the high management. |
| 10 | Demonstrate the progress via indicators to the high management to increase the interest in the energy management. |
| 11 | Training, motivating, and providing information to the employees of the company about energy management activities. |

The five stages of the IEnMS (Industrial Energy Management System) framework are explained in a simplified manner. Considering a company wants to implement an energy management system to reduce greenhouse gas emissions, save energy, and lower its energy costs, it can employ IoT applications for its Energy Management System (EnMS). To begin with, the company has to familiarize itself with IoT components to understand what the IoT is and how it functions. Next, IoT platforms and their key features have to be studied. Although the company may require an IoT platform for its business application, it may be uncertain which platforms offer the required, specific features, and which is the most suitable one for the company's purposes, as explained in Publication I. Once the company gains knowledge about the IoT, IoT platform factors, and the features provided by those platforms, it proceeds to stage 1.

In stage 1, the company evaluates each of the 21 IoT platform factors highlighted as important in selecting a platform as presented in Publication I. This evaluation assists the company in formulating its business requirements. In stage 2, these factors are categorized as either "required" (R) or "important" (I) for its business needs. For instance, the company may discover that its required considerations (R) include aspects like scalability, stability, system performance, user-friendly interface, edge intelligence, time to market, flexibility, and prior experience. The important considerations (I) can include factors like pricing, security, data analytics, disaster recovery, and interoperability.

Moving on to stage 3, the company compares the R and I factors with the identified features of IoT platforms. The company can select a platform that meets both its I and R requirements. If multiple IoT platforms fully satisfy the company's I and R requirements, it can choose one that excels in meeting its I requirements. Occasionally, none of the IoT platforms may fulfill all the I and R requirements. In such cases, the company can explore alternative IoT platforms for its IoT application. In the case of the IEnMS, only the methodology of how to select an IoT platform based on the company's requirements is provided.

### 3.3.4 Publication VI: Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing.

This research focuses on the implementation of the IoT and big data in smart grids by answering the following research question:

***How smart grid data are processed using the IoT, big data, and cloud and edge computing?***

Power electronics and internet and communication technologies (ICT) have made great strides in recent years, enabling the transition from the traditional electric grid of the 20th century to the modern, smart electric grids of the 21st century. Additionally, modern smart grid (SG) operations are impacted by new technological paradigms like the Internet of Things (IoT) by enhancing communication, fostering stronger client interactions, and managing the enormous amount of data generated by smart devices. In this regard, IoT solutions for data collection, connectivity, and intelligent analytics are becoming more

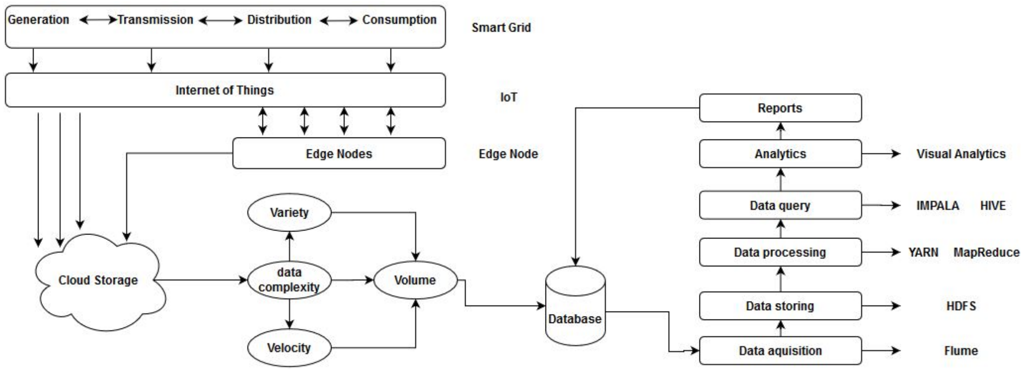



Figure 3.11: IoT and edge computing in smart grids using big data analytics (Publication VI).

prevalent in SG applications.

Cloud computing is frequently used in smart grid applications to handle the computational needs for effective electricity supply. The number of devices connected to the smart grid has recently expanded as a result of the emergence of the IoT, including consumer electronics, measuring tools, and electrical and electronic equipment like smart power converters, phasor measurement units, and smart meters. A significant amount of structured, semi-structured, and unstructured data is produced by these heterogeneous devices that are present in all four stages of a smart grid: generation, transmission, distribution, and consumption. Utilizing cloud computing to collect, store, and process such massive amounts of data results in issues with bandwidth, latency, disaster recovery, and cost.

With the increase in data from devices in an industrial setting leads to a number of issues. At the beginning, a huge amount of data is moved to the cloud for analysis, but the majority of the data may not be useful to the operations. As a result, this data transmission leads to increased traffic to a central repository and higher storage costs. Second, significant delays can occur when sending important data with a low latency of a few seconds to a few milliseconds, which is necessary for critical industrial operations. Third, both uploading and downloading data from the cloud can be very expensive.

Edge computing is used to overcome the aforementioned problems in cloud computing. The benefits of edge computing are that it moves data analysis and services away from centralized servers, and a lot of data analysis is performed at the source of data collection.

As indicated in Figure 3.11, an architecture that uses the IoT and edge nodes to obtain SG data was developed for this work. This is the process by which SG data are created, processed, and analyzed. Numerous smart meters, sensors, and digital devices produce the data throughout a certain time period. Generation facilities (e.g., wind farms, solar farms, conventional power plants), transmission and distribution networks (e.g., phasor measuring units), or clients (e.g., residential homes, electric vehicles, commercial buildings, industries) can all generate data. Data from the environment, such as weather, humidity, temperature, and pressure information, can also be collected. Social media can be used to gather some useful data, such as details on outside events. Data produced from numerous sources improve the grid reliability. IoT devices like sensors and actuators transfer the generated data to the IoT network by employing the network technologies 3G/4G/5G, ZigBee, Wi-Fi, Bluetooth, and wired connectivity.

As the edge nodes are closer to the locations where data are collected in comparison with the cloud servers, they are usually used to process the critical data that need to be handled rapidly and without delay. The edge network should be constructed with consideration of elements including task complexity, server processing power, and network architecture in order to effectively meet the latency requirements. Edge computing reduces the quantity of data that need to be transferred to the cloud by dispersing the data among various edge servers for computational workload. This has an impact on the bandwidth. When the total number of jobs surpasses the combined processing power of the edge servers, the cloud and edge computing work together to deliver a high bandwidth as the bits are transported to the cloud server. A further advantage of edge computing in an SG is the reduction in failure rate. The edge computing services of the other sections of the grid will function regularly, without any issues, even if there is an electrical outage in a specific sector of the grid. On the other hand, if all of the grid's processing is done in the cloud and there is a power supply failure because of a natural disaster in the cloud infrastructure, the entire network will go down.

# 4 Conclusions

The dissertation is related to the digitalization of industries in the era of the IR4.0. In the work, the main focus was on the industrial digitalization employing the latest technologies like the IoT, IoT platforms, big data, and edge computing. A key question that arises during the digitalization of different industrial processes is: **Is it possible to create a unified framework for the digitalization of industrial sectors based on the IoT in integration with technologies like big data and analytics and edge computing?** This dissertation contributed to solve this challenge by answering the following more specific questions:

**RQ 1:** What would a unified framework look like that can be used for the selection of an IoT platform based on companies' business requirements?

**RQ 2:** What happens when such a unified framework is applied to a specific domain of Industrial Energy Management?

**RQ 3:** Can such an approach be deployed in different industrial sectors?

To answer the above-mentioned main question and subquestions, detailed research was conducted and documented in the form of Publications I–VI. Various research methods were employed to address the importance and challenges in using the latest technologies like the IoT, IoT platforms, big data, and edge computing in industrial processes from different perspectives to answer the research questions. During the IR4.0 there is immense growth in the IoT devices, IoT platforms, and their applications.

Therefore, digitalization of industrial processes and services, reliable and efficient IoT platforms, efficient data analytics, and fast and reliable networking are crucial for successfully deploying and adopting these technologies. Based on the company's business requirements and the features provided by the IoT platforms, a theoretical five-stage framework for IoT platform selection was developed, which provides a solid base for the companies that want to automate their business processes (shifting their business processes to the cloud) by using the latest technologies like the IoT, big data, and edge computing. However, the adoption and implementation of these technologies may be challenging, and companies may hesitate over shifting their business to the cloud. This may be due to the lack of experience in the IoT and IoT platforms, and hesitation to select a suitable IoT platform among hundreds of IoT platforms on the market for their business processes.

To overcome the above-mentioned challenges, the dissertation focused first on explaining the building blocks of the IoT that work together to accomplish a specific task, and on highlighting the key factors of an IoT platform. Second, a five-stage conceptual framework was designed for selecting a suitable IoT platform for industrial business application based on the business requirements. Third, different frameworks for implementing the IoT in different fields of the industrial sector were designed to improve the data collection and facilitate the business processes.

The objectives of the dissertation were achieved through the studies presented in Publications I–VI as follows.

**Publication I:** The main components of IoT technology that work together to achieve a specific goal were first identified, and then, 21 key factors of an IoT platform were presented and studied. The top five IoT platforms (Amazon web services, Microsoft azure, Google Cloud, IBM Watson, and Oracle IoT) based on their market shares were studied. Based on the findings, a five-stage theoretical framework was designed for the selection of a suitable IoT platform for industrial business needs. To the best of the author's knowledge, no previous studies have addressed the 21 key factors of an IoT platform in a systematic manner, and so far, there has been no framework designed for the selection of an IoT platform for a company's business application. This study provides a road map for companies that wish to shift their business to the cloud and need information about the IoT, its components, IoT platforms, and their key components to select a suitable IoT platform for their business application. The study was limited to top five IoT platforms based on their market shares. The same study can be used in various industrial applications to improve various business processes and plans for future business decisions; this is achieved by capturing data using IoT sensors, and processing and storing the data by using an IoT platform.

**Publication II:** The focus was on industrial energy management motivated by the fact that a considerable amount of the world's energy consumption is due to industrial activities, which greatly increase greenhouse gas emissions. Implementing an energy management system is one of the best methods to cut down on the energy use in the industrial sector, but thus far, there has not been a comprehensive framework for an IEnMS that uses the Internet of Things (IoT) and big data. To solve this problem, an architecture for the IEnMS was designed that can be used by the industries to reduce their energy use, greenhouse gas emissions, and energy costs. The task was accomplished by utilizing the work done in Publication I, in which the building blocks of the IoT were defined. In Publication II, the IoT, big data and analytics were used as the modern technologies to achieve the requirements of the IEnMS. Later in Publication V, the five-stage IoT platform framework developed in Publication I was used for selecting an IoT platform for the IEnMS. Specifically, an IoT application was employed to collect data by using sensors and actuators, and the collected data were sent to big data for analytics. The IoT platform was then used for developing an IoT-based IEnMS.

**Publication III:** A P2X process was used to convert renewable energy into storable hydrogen, chemicals, and fuels through electrolysis and subsequent synthesis with $CO_2$. The study highlighted the main contributions of Publication I, in which the main six building blocks of the IoT were defined. The study also highlighted the contributions provided by Publication II, in which the big data and analytics were explained, and the operation and requirements of the IoT and big data were elucidated; these contributions include the use of advanced data-driven methods and latest technologies like the Internet of Things (IoT), big data analytics, and machine learning for the efficient operation of P2X cogeneration plants. The findings were summarized into different operation architectures that use different sensors to collect data from different sources and store them in big data for data analytics and future predictions by using machine learning.

**Publication IV:** The proposed framework designed in Publication I for the selection of an IoT platform for industrial application based on the company's business requirements was employed in a pyrolysis process, which is used to efficiently dispose of the accumulated carbon fiber-reinforced polymer (CFRP) composite wastes. In the study, a pyrolysis process was used to recycle CFRP composite waste, and a theoretical five-stage framework was employed for the selection of a suitable IoT platform to be employed to manage the process as a CPS. The suggested framework consists of sensors and actuators that are used to gather data and communicate with a centralized management. It was designed as a platform that will articulate and manipulate data to meet the needs of the recycling process, which was modeled computationally through logical relationships between physical entities. The publication showed the benefits of using the IoT platform selection framework in the industrial sector.

**Publication V:** The focus was on the application of the framework introduced in Publication I into the IEnMS studied in Publication II, where the latest technologies like the IoT and big data were used for the proper utilization of the IEnMS in an industrial process to save energy and reduce the energy cost. As there are hundreds of IoT platforms on the market, and thus, selecting a suitable IoT platform that can meet the business requirements is very difficult, the proposed five-stage theoretical framework was used to select a suitable IoT platform for the IEnMS. The business requirements of the IEnMS were identified and compared with the features provided by the IoT platform. The platform that best suits the requirements based on the proposed framework was selected. This approach provides an objective methodology that can be used to select the most suitable IoT platform for different IEnMSs based on their particular requirements.

**Publication VI:** A theoretical framework was used for the effective collection of data from a large number of devices linked to a smart grid as well as for the processing, storage, and display of the gathered data using a combination of the IoT, edge computing, big data, and analytics. The publication provided in-depth discussion of the several advantages of edge computing and big data, including latency, bandwidth, disaster recovery, and cost in the SG system as a whole, from data collecting to data presentation. The publication elaborated on some benefits of edge computing and big data related to, e.g., latency, bandwidth, disaster recovery, and price of the entire SG system, starting from data collection to data visualization. The study provided a theoretical discussion of improvement suggestions for the smart grid.

These results clearly indicated that it is indeed possible to propose a unified framework to support different industrial sectors to move toward digitalization of their processes, while paying special attention to aspects related to energy and sustainability. In the following, the dissertation is concluded with a summary of the key contributions and potential future research directions.

**Summary:** The dissertation studied and explained the IoT, IoT platforms, and their key factors in order to answer the question of how to implement the latest technologies in the automation processes of industrial applications. The contributions of this work can be summarized as follows:

- Definition and explanation of the IoT and its main components.
- Highlighting of 21 key factors of an IoT platform.
- Design of a theoretical framework for the selection of an IoT platform.
- Proposal for a theoretical framework for the IEnMS using the IoT, big data, and data analytics to construct an effective cyber-physical system architecture including steps from data acquisition to the end-user decision-making process.
- Design of a working architecture for the P2X process for long-term storage of renewable energy using the latest technologies like the IoT, big data analytics, and machine learning.
- Theoretical implementation of an IoT-based framework into the pyrolysis process to recycle CFRP composite waste to manage the process based on the principles of cyber-physical systems.
- Demonstration of the benefits of using the IoT, edge computing, and big data in different industrial applications.

**Future work:** These contributions indicate different directions for further development. This dissertation provides a theoretical framework and is limited to only five IoT platforms (AWS, Microsoft Azure, Google cloud IoT, IBM Watson IoT, Oracle IoT) selected based on their market shares. In the questionnaire survey conducted in the study, the questions presented to the industry experts were limited to certain specific points, and there were no open questions where the respondents could have added some more factors that they considered important. In the context of the survey, the threats to validity refer to the different sources of threats that can affect the reliability and accuracy of the results of the survey. The validity of the conclusions drawn from the survey can be affected by the threats. Those threats can be, for instance, selection of a wrong population sample, nonresponse, social relations, and wrong questions asked. To overcome those threats, the survey has to be designed carefully, the questions can be tested in advance, and random sampling can be selected, with the target of reaching a high response rate.

In this study, two surveys were conducted, one survey about the key factors of an IoT platform (questions in Table 3.2 of Publication I) and another one about the use of present technologies in the IEnMS (questions in Table 3.4 of Publication II). In both the surveys, the questionnaires were designed in such a way that the questions were specifically related to the main topics, they were kept simple, and presented in a logical order. The population sample selected for the survey consisted of experts on the topics. The sample size was small (15) in the first survey; however, in Delphi studies the minimum number of respondents can be as low as three, and the maximum number can be even eighty [45]. The survey had two rounds, and IoT experts were selected as the respondents. Fourteen out of the fifteen candidates responded to the questionnaire, and thus, the response rate was high.

The second survey was about the use of modern technologies in the IEnMS, and experts in large companies were selected for the survey. Experts from ten companies replied to the survey, and thus, the response rate was somewhat low. The results of the surveys show a high interest of the respondents in the survey questions, and most of the respondents agreed on the statements presented in the questions. From our viewpoint, the results of our surveys are accurate and reliable.

In the future, the same framework can be applied to any other IoT platform. In the future, a practical approach can be applied to develop business applications for a specific industry by selecting a suitable IoT platform (based on the business requirements of that industry) using the proposed five-stage IoT platform selection framework. For example, an automatic tool could be developed for selecting the most suitable IoT platform given a set of requirements. Another option is to empirically assess the performance of the IEnMS in different companies. In addition, the deployment of testbeds of industrial cyber-physical systems following the framework proposed here may lead to new research challenges.

# References


[1] N. Jazdi, "Cyber physical systems in the context of industry 4.0," in *2014 IEEE international conference on automation, quality and testing, robotics*. IEEE, 2014, pp. 1–4.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE communications surveys & tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[3] C. Greer, M. Burns, D. Wollman, and E. Griffor, "Cyber-physical systems and internet of things," *National Institute of Standards and Technology*, pp. 1900–202, 2019. [Online]. Available: https://doi.org/10.6028/NIST.SP.1900-202

[4] X. Jia, Q. Feng, T. Fan, and Q. Lei, "Rfid technology and its applications in internet of things (iot)," in *2012 2nd international conference on consumer electronics, communications and networks (CECNet)*. IEEE, 2012, pp. 1282–1285.

[5] D. G. Pivoto, L. F. de Almeida, R. da Rosa Righi, J. J. Rodrigues, A. B. Lugli, and A. M. Alberti, "Cyber-physical systems architectures for industrial internet of things applications in industry 4.0: A literature review," *Journal of manufacturing systems*, vol. 58, pp. 176–192, 2021.

[6] A. Mosenia and N. K. Jha, "A comprehensive study of security of internet-of-things," *IEEE Transactions on emerging topics in computing*, vol. 5, no. 4, pp. 586–602, 2016.

[7] P. H. Nardelli, *Cyber-physical systems: Theory, methodology, and applications*. John Wiley & Sons, 2022.

[8] H. Hejazi, H. Rajab, T. Cinkler, and L. Lengyel, "Survey of platforms for massive iot," in *2018 IEEE international conference on future IoT technologies (future IoT)*. IEEE, 2018, pp. 1–8.

[9] L. Georgios, S. Kerstin, and A. Theofylaktos, "Internet of things in the context of industry 4.0: An overview," *International Journal of Entrepreneurial Knowledge*, 2019.

[10] M. Gigli, S. G. Koo *et al.*, "Internet of things: services and applications categorization." *Adv. Internet Things*, vol. 1, no. 2, pp. 27–31, 2011.

[11] F. H. Mohammed and R. Esmail, "Survey on iot services: Classifications and applications," *International Journal of Science and Research*, vol. 4, no. 1, pp. 2124–2127, 2015.

[12] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, "A gap analysis of internet-of-things platforms," *Computer Communications*, vol. 89, pp. 5–16, 2016.

[13] J. Guth, U. Breitenbücher, M. Falkenthal, F. Leymann, and L. Reinfurt, "Comparison of iot platform architectures: A field study based on a reference architecture," in *2016 Cloudification of the Internet of Things (CIoT)*. IEEE, 2016, pp. 1–6.

[14] S. Narula, A. Jain *et al.*, "Cloud computing security: Amazon web service," in *2015 Fifth International Conference on Advanced Computing & Communication Technologies*. IEEE, 2015, pp. 501–505.

[15] V. Gazis, M. Görtz, M. Huber, A. Leonardi, K. Mathioudakis, A. Wiesmaier, F. Zeiger, and E. Vasilomanolakis, "A survey of technologies for the internet of things," in *2015 international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2015, pp. 1090–1095.

[16] I. . i Scoop, "Industry 4.0 and the fourth industrial revolution explained," 2023, https://www.i-scoop.eu/industry-4-0/ [Accessed: (22nd November 2023)].

[17] A. Bastos, M. L. S. C. De Andrade, R. T. Yoshino, and M. M. D. Santos, "Industry 4.0 readiness assessment method based on rami 4.0 standards," *IEEE Access*, vol. 9, pp. 119 778–119 799, 2021.

[18] J. H. Nord, A. Koohang, and J. Paliszkiewicz, "The internet of things: Review and theoretical framework," *Expert Systems with Applications*, vol. 133, pp. 97–108, 2019.

[19] J. K. Gerrikagoitia, G. Unamuno, E. Urkia, and A. Serna, "Digital manufacturing platforms in the industry 4.0 from private and public perspectives," *Applied Sciences*, vol. 9, no. 14, p. 2934, 2019.

[20] I. T. Christou, N. Kefalakis, J. K. Soldatos, and A.-M. Despotopoulou, "End-to-end industrial iot platform for quality 4.0 applications," *Computers in Industry*, vol. 137, p. 103591, 2022.

[21] F. Hoffman, "Industrial internet of things vulnerabilities and threats: What stakeholders need to consider." *Issues in Information Systems*, vol. 20, no. 1, 2019.

[22] M. J. Perry, *Evaluating and Choosing an IoT platform*. O'Reilly Media, 2016.

[23] Igor Ilunin, "How to Choose Your IoT Platform-Should You Go Open-Source?" 2017. [Online]. Available: https://medium.com/iotforall/ how-to-choose-your-iot-platform

[24] C. McClelland, "Leverege—how to choose the best iot platform," 2018, accessed: (2023-05-19). [Online]. Available: https://www.leverege.com/iot-ebook/iot-platforms-explained

[25] M. Y. Santos, J. O. e Sá, C. Andrade, F. V. Lima, E. Costa, C. Costa, B. Martinho, and J. Galvão, "A big data system supporting bosch braga industry 4.0 strategy," *International Journal of Information Management*, vol. 37, no. 6, pp. 750–760, 2017.

[26] S. Sagiroglu, R. Terzi, Y. Canbay, and I. Colak, "Big data issues in smart grid systems," in *2016 IEEE international conference on renewable energy research and applications (ICRERA)*. IEEE, 2016, pp. 1007–1012.

[27] B. P. Bhattarai, S. Paudyal, Y. Luo, M. Mohanpurkar, K. Cheung, R. Tonkoski, R. Hovsapian, K. S. Myers, R. Zhang, P. Zhao *et al.*, "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions," *IET Smart Grid*, vol. 2, no. 2, pp. 141–154, 2019.

[28] M. M. Nair, A. K. Tyagi, and N. Sreenath, "The future with industry 4.0 at the core of society 5.0: Open issues, future opportunities and challenges," in *2021 international conference on computer communication and informatics (ICCCI)*. IEEE, 2021, pp. 1–7.

[29] J. Waldo, G. Wyant, A. Wollrath, and S. Kendall, "A note on distributed computing," in *International Workshop on Mobile Object Systems*. Springer, 1996, pp. 49–64.

[30] S. Rani, A. Kataria, and M. Chauhan, "Fog computing in industry 4.0: Applications and challenges—a research roadmap," *Energy conservation solutions for fog-edge computing paradigms*, pp. 173–190, 2022.

[31] Y. Ai, M. Peng, and K. Zhang, "Edge computing technologies for internet of things: a primer," *Digital Communications and Networks*, vol. 4, no. 2, pp. 77–86, 2018.

[32] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[33] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE internet of things journal*, vol. 3, no. 5, pp. 637–646, 2016.

[34] S. Yi, Z. Hao, Z. Qin, and Q. Li, "Fog computing: Platform and applications," in *2015 Third IEEE workshop on hot topics in web systems and technologies (HotWeb)*. IEEE, 2015, pp. 73–78.

[35] K. Ha, Z. Chen, W. Hu, W. Richter, P. Pillai, and M. Satyanarayanan, "Towards wearable cognitive assistance," in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, 2014, pp. 68–81.

[36] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, and A. Ahmed, "Edge computing: A survey," *Future Generation Computer Systems*, vol. 97, pp. 219–235, 2019.

[37] E. R. Griffor, C. Greer, D. A. Wollman, and M. J. Burns, "Framework for cyber-physical systems: Volume 1, overview," 2017.

[38] R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, *Cyber-physical systems: the next computing revolution*. ACM digital library, 2010.

[39] K. Cao, S. Hu, Y. Shi, A. W. Colombo, S. Karnouskos, and X. Li, "A survey on edge and edge-cloud computing assisted cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7806–7819, 2021.

[40] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *2011 international conference on wireless communications and signal processing (WCSP)*. IEEE, 2011, pp. 1–6.

[41] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *2008 The 28th International Conference on Distributed Computing Systems Workshops*. IEEE, 2008, pp. 495–500.

[42] Y. Tan, M. C. Vuran, and S. Goddard, "Spatio-temporal event model for cyber-physical systems," in *2009 29th IEEE International Conference on Distributed Computing Systems Workshops*. IEEE, 2009, pp. 44–50.

[43] T. Sanislav and L. Miclea, "Cyber-physical systems-concept, challenges and research areas," *Journal of Control Engineering and Applied Informatics*, vol. 14, no. 2, pp. 28–33, 2012.

[44] M. Ullah, P. H. Nardelli, A. Wolff, and K. Smolander, "Twenty-one key factors to choose an iot platform: Theoretical framework and its applications," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10 111–10 119, 2020.

[45] E. Ogbeifun, J. Agwa-Ejon, C. Mbohwa, and J. Pretorius, "The Delphi technique: A credible research methodology," 2016. [Online]. Available: https://hdl.handle.net/10210/92867

# Publication I

Ullah, M., Nardelli, P. H. J., Wolff, A., and Smolander, K.

**Twenty-One Key Factors to Choose an IoT Platform: Theoretical Framework and Its Applications**

Reprinted with permission from
*Internet of Things Journal*
vol. 7, no. 10, pp. 10111–10119, Oct. 2020, doi: 10.1109/ JIOT.2020.3000056.

# Twenty-One Key Factors to Choose an IoT Platform: Theoretical Framework and Its Applications

Mehar Ullah, *Member, IEEE*, Pedro H. J. Nardelli, *Senior Member, IEEE*, Annika Wolff, and Kari Smolander

***Abstract*—Internet of Things (IoT) refers to the interconnection of physical objects via the Internet. It utilizes complex back-end systems that need different capabilities depending on the requirements of the system. IoT has already been used in various applications, such as agriculture, smart home, health, automobiles, and smart grids. There are many IoT platforms, each of them capable of providing specific services for such applications. Finding the best match between application and platform is, however, a hard task as it is difficult to understand the implications of small differences between platforms. This article builds on previous work that has identified 21 important factors of an IoT platform, which were verified by the Delphi method. We demonstrate here how these factors can be used to discriminate between five well-known IoT platforms, which are arbitrarily chosen based on their market share. These results illustrate how the proposed approach provides an objective methodology that can be used to select the most suitable IoT platform for different business applications based on their particular requirements.**

***Index Terms*—Components, features, key factors, platform.**

## I. INTRODUCTION

THE Internet-of-Things (IoT) concept was first coined by Keven Ashton in 1999 during a presentation to Proctor and Gamble and later referenced by him in the MIT Auto-ID Center [2]. IoT is one of the fastest growing technologies that is gaining momentum in various domains, such as transportation, healthcare, industrial automation, education sector, etc. The main idea of IoT is to connect the physical world with the digital world [3]. The foundation technology for IoT is the RFID technology, which is used to identify, track, and monitor any object with RFID tags and allow microchips to transmit the identification information to a reader through wireless communication [4]. Nowadays, IoT applications have already moved further away than just simple RFIDs, incorporating different sources of data collection from sensors. This data stream needs to be moved somewhere where this (big) data can be processed using, for example, machine learning techniques. This place is what we call an *IoT platform*.

For companies to run their specific IoT applications, an IoT platform is then needed. The IoT platform provides important services and features to applications: endpoint management, connectivity and network management, analysis and processing, data management, application development, security, event processing, monitoring, access control, and interfacing [5]. From 2015 onward, there has been a rapid growth in IoT technologies so that the number of connected devices and platforms has steadily increased. For example, there were 260 IoT platforms in 2015, which increased to 360 in 2016 and 450 in 2017 [6]. Due to technological improvements, new IoT devices emerged and the requirements of the IoT applications and platforms changed [7]. Such a technological change creates many challenges for businesses, governments, and companies, which have little experience with the infrastructure of IoT and IoT platforms. Selecting a suitable IoT platform among all existing options is a tricky task since this decision needs to incorporate not only the current needs but also the potential future ones [8].

There are hundreds of IoT platforms in the market, most with similar functionality with differences related to their implementation and underlying technologies [8]. Our aim with this article is to first highlight the key building blocks of IoT for the understanding of functionality and significance of IoT and identify and verify the key factors of an IoT platform. With the key factors in hand, we could then propose an objective and general methodology to compare the different service providers. In this case, this article develops a theoretical framework that will support companies in selecting a suitable IoT platform for their business needs. To carry out this article, we follow three steps: 1) data collection; 2) data verification and characterization; and 3) application of the proposed framework. These are the questions that were used to guide this article: 1) what is IoT as well as its building blocks? 2) what are the important factors of an IoT platform? and 3) what factors should be considered for selecting an appropriate IoT platform for specific organizations?

This article covers and extends our preliminary work [1] in which we have identified 21 IoT platform factors from the literature and verified those factors using the Delphi method. This contribution builds upon [1] aiming at developing a

Manuscript received November 1, 2019; revised February 10, 2020 and April 10, 2020; accepted May 29, 2020. Date of publication June 8, 2020; date of current version October 9, 2020. This work was supported in part by the Academy of Finland through ee-IoT under Grant 319009, FIREMAN Consortium under Grant CHIST-ERA/326270, and EnergyNet Fellowship under Grant 321265 and Grant 328869. This article was presented in part at the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, May 2019. *(Corresponding author: Mehar Ullah.)*

Mehar Ullah and Pedro H. J. Nardelli are with the Department of Electrical Engineering, LUT University, 53850 Lappeenranta, Finland (e-mail: mehar.ullah@lut.fi).

Annika Wolff and Kari Smolander are with the Department of Software Engineering, LUT University, 53850 Lappeenranta, Finland.

Digital Object Identifier 10.1109/JIOT.2020.3000056

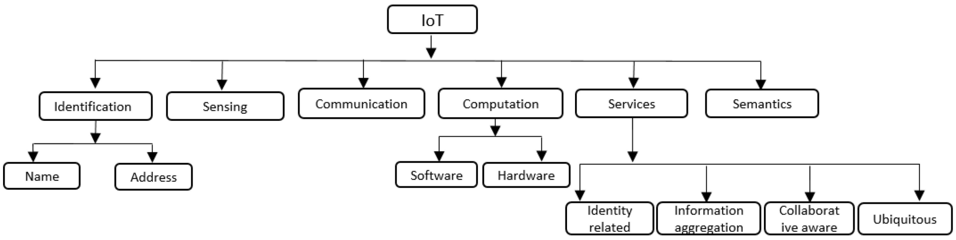



Fig. 1. IoT building blocks, following the structure proposed by [9].

theoretical framework by comparing the 21 key IoT platform factors with the features provided by the five popular IoT platforms. Our goal is to provide an objective while general methodology that different organizations can apply when selecting the most suitable platform based on their particular needs. In other words, this article will support such organizations to carry out a detailed analysis of their own requirements and understand the key IoT platform factors in order to find the best match.

The remainder of this article is organized as follows. In Section II, we review the IoT building blocks. In Section III, we identify the key factors of an IoT platform. Section IV introduces the research method including the results of the Delphi study. Section V focuses on the comparison of the main IoT platforms. Section VI concludes this article indicating potential research paths.

## II. IoT Building Blocks

To understand the functionality and significance of IoT, it is essential to understand its building blocks; they are the components of IoT, which work together to deliver its functionality. There are six IoT building blocks that work together and provide functionality [3], as shown in Fig. 1. In the following, we will explain each of them in more detail.

*Identification Block:* The identification method is used to identify devices in the network. Devices are identified with the object ID, which is the name of the device and the object address, which provides the address of the device in the communication network [10]. The main addressing methods of IoT objects are IPv6 and IPv4 [3].

*Sensing Block:* Sensors are used for collecting the data of objects/environment in the communication network and sending the collected data to the destination database or to the cloud. The data collected is analyzed in the cloud. Actuators, i.e., hardware mechanical devices such as switches, are also used in IoT platforms and operate in the opposite way to a sensor [3], [11], [12].

*Communication Block:* It contains many heterogeneous objects that exchange data and various services with each other and with the platform. The communication block contains IoT communication protocols, such as MQTT and CoAP that are used to connect different objects to IoT and to send data from those connected objects to the management system. The sensors and other devices are connected to the Internet by communication technologies, such as ZigBee, NFC, UWB, Wi-Fi, SigFox, and BLE [3], [7].

*Computation Block:* The computation block consists of two parts: 1) hardware and 2) software. Many hardware platforms have been built to run IoT applications, for example, Intel Galileo, Raspberry PI, Gadgeteer, UDOO, and Arduino. Similarly, there are many software platforms that are used to perform the functionalities of IoT. The main software platform is the operating system that runs throughout almost the whole activation time of the device. The cloud platform is also a computational component of the IoT; it enables small objects to send data to the cloud and it facilitates big data processing in real time and helps the end user to obtain knowledge extracted from the big data [3], [7].

*Services Block:* IoT services aid IoT application developers by providing a starting point for development. When developers know the services available, they mainly focus on building the application rather than designing the service and architecture for supporting the IoT application. IoT services are divided into four categories. Identity-related services can be divided into two categories: 1) active and 2) passive. Services that broadcast information and have a constant power or take power from the battery are active identity-related services. Active identity-related services can transmit or send information to another device. Passive identity-related services have no power source and need some external device or mechanism to transmit its identity. Passive identity-related services can only read information from devices. Information aggregation services refer to the actions of collecting data from sensors, processing that data, and transferring it to the IoT application for processing. Collaborative-aware services use the data provided by the information aggregation services to make decisions and react accordingly. Ubiquitous services provide collaborative-aware services anytime to anyone who needs it anywhere [11], [13], [14].

*Semantic Block:* IoT provides different services, for which it needs knowledge, and in order to get that knowledge in an effective way, IoT uses different machines. Knowledge extraction can include finding and using resources, modeling information, and recognizing and analyzing data to reach some decision and provide the correct service. So it can be claimed that the semantic block is the brain of the IoT [3], [7], [11].

## III. Key Factors of IoT Platform

An IoT platform is the main part of an IoT solution. There are hundreds of IoT platform vendors in the market, and finding and selecting a suitable IoT platform that is reliable and scalable is difficult. However, consideration of some key factors prior to making a platform selection decision can enable companies to find and select an appropriate IoT platform for their business. The platform requirements are context specific and it is not necessary that a platform includes all the factors discussed below, but can have a maximum. These factors were identified from literature by studying various IoT platforms [15]–[19], articles [7], [20], [21], and websites, such as [22]–[24].

*Stability:* There are hundreds of platforms in the market, which might have some open issues. Some platforms might fail to deliver services to clients. Thus, a platform should be chosen that has high chance of survival in the market. Information about the platform can be obtained from previous customers using the same platform [7].

*Scalability and Flexibility:* Initially, a company might be small and operates in a small business area but, ideally, over time, the business will expand and with this growth, the business area will also expand. Thus, to ensure that the IoT platform can support the business throughout its development, the platform should be scalable to business needs [7]. Similarly, the platform should be flexible with regard to technology since modern technology and market demand change rapidly.

*Pricing Model and Business Case:* Some platform providers offer a low price for a period at the start of a contract agreement, after which the price increases greatly. Additionally, some providers offer a low price to attract customers, but the contract includes limited features and additional features have a significant cost if included. Thus, a platform should be selected that offers full features for the business at a cost that suits the company's budget [7], [25].

*Security:* It is an important aspect of IoT that all platforms should have with high quality. The security may be device-to-cloud network security, data encryption, application authentication, secure session initiation, application authentication, cloud security, and device security (authentication and up-to-date certification) [26].

*Time to Market:* When selecting an IoT platform, the questions of time to market and how the platform provider will support the business during the journey from product conception to sale should be considered. Some IoT platform providers offer quick-start packages for new customers, which can speed up product development, reduce time to market, and offer better solutions [26].

*Data Analytics and Visualization Tools:* Before selecting an IoT platform, prospective IoT platform users should establish which platform offers the best capabilities to aggregate, analyze, and visualize data. In particular, users should consider how the IoT platform integrates leading analytics toolsets and uses them to replace built-in functionality. Data analysis and information visualization requirements should be identified before selecting an IoT platform [13].

*Data Ownership:* A complicated issue with IoT data is the ownership of the data. Different jurisdictions have different laws and legal interpretations. For example, the European Union (EU) has different rules and regulations regarding data ownership than the United States (U.S.) [27]. Therefore, it is important to have knowledge of data rights and the territorial scope of data protection for the IoT platform provider.

*Ownership of Cloud Infrastructure:* The hardware infrastructure layer is expensive and some smaller IoT platform providers only offer the software layer. Some providers certify their platform on single or multiple leading public cloud providers and mostly run their services on a single leading platform. The compatibility of the broader enterprise cloud with the IoT platform provider should be checked [28].

*Extent of Legacy Architecture:* The connectivity in an existing IoT is often unknown, and IoT devices are designed to work with a variety of infrastructure systems. Therefore, when selecting an IoT platform, businesses and organizations should ascertain how new generations of technology can interlock with older technology [29].

*Protocol:* The important protocols supported by IoT platforms are MQTT, HTTP, AMQP, and CoAP. Due to its binary nature, MQTT is extremely lightweight and has much lower overheads. As a result of the development in technology, new devices are coming onto the market. The selected IoT platform should support new protocols and enable the easy upgrade of these protocols [23], [30].

*System Performance:* In an IoT platform, when an event happens, a rule-based trigger might be invoked automatically. Since they support such a method, as larger numbers of devices connect to the IoT platform, the average time to analyze and handle each event increases. Prior to the selection of an IoT platform, it should be noted what steps the provider has taken to maintain the IoT platform performance high enough [31].

*Interoperability:* The IoT platform solution is middleware. The data collected will be used by many applications and may not be available on the platform itself. Consequently, the selected IoT platform should support integration with open-source ecosystems. Interoperability will enable the organization to gain higher productivity [32], [33].

*Redundancy and Disaster Recovery:* Problems sometimes occur in the IT infrastructure, either natural or man made, and IoT platform providers should have dedicated infrastructure to handle data during such occurrences. Issues that require consideration include the data backup plan schedule and whether the IoT platform has failover cluster provision [23].

*Attractive Interface:* The interface provided by the IoT platform should be simple, attractive, and user friendly so that it is easy for customers to use its functionalities. All the services offered to the customers should be easy to access.

*Application Environment:* Three aspects of the application environment should be considered before selecting an IoT platform: which applications are available out of the box, what are the characteristics of the application development environment, and what are the common interfaces [28].

*Hybrid Cloud:* Some IoT platforms can fit with existing IT systems hosted on company premises. In such situations, a hybrid cloud is very useful as mission-critical or business-sensitive processes can be handled locally, while public and less critical operations can be managed by the platform [23].

*Platform Migration:* Over time, and as the company grows, the IoT platform may be unable to meet all the company's requirements. Thus, a bigger IoT platform provider may

TABLE I
BASIC FEATURES PROVIDED BY THE FIVE IoT PLATFORMS

| Features | AWS | Microsoft Azure | Google Cloud IoT | IBM Watson IoT | Oracle IoT |
|---|---|---|---|---|---|
| **Security** | Link Encryption (TLS), Authentication (Sig V4, X.509) | Link Encryption (SSL/TSL) | SSL/TLS | Link Encryption (TLS), Authentication (IBM cloud SSO), Identity Management (LDAP) | REST API |
| **Data analytics** | Real Time analytics (Rule engine, Kinesis, AWS Lambda) | Real Time analytics | Real Time analytics (Cloud IoT Core) | Real Time analytics (IBM IoT Real time insights) | Real Time analytics |
| **Protocols** | MQTT, HTTP1.1 | MQTT, HTTP, AMQP | MQTT | MQTT, HTTPS | MQTT, HTTP |
| **Visualization tool** | AWS IoT dashboard | web portal | Google data studio (Dashboard) | web portal | web portal |
| **Data format** | JSON | JSON | JSON | JSON, CSV | CSV, REST API |
| **Application Environment** | Java, C, NodeJs, Javascript, Python, SDK for Arduino, iOS, Android | .Net, UWP, Jave, C, NodeJS, Ruby, Android,iOS | Go,Java, Python, .NET, NodeJS, php, Ruby | C#, C, Python, Java, NodeJS | Java, iOS, Javascript, C, Android |

be needed. Consequently, companies should ensure that the selected IoT platform provider provides clearly documented interfaces, schema, and API for any possible future migration to other IoT platforms [7], [23].

*Previous Experience:* Prior to selection, a company should check whether the IoT platform provider has some experience of work similar to that of the company application. Successful working experience in the same area can be considered a good sign [26].

*Bandwidth:* For the efficient movement of information and communication between the processing components, the IoT platform needs low-latency and high bandwidth networking. Thus, it should be ascertained that a potential IoT platform provider has a large data pipe and that there is sufficient room to grow [23], [34].

*Edge Intelligence and Control:* The future of IoT platforms is moving toward distributed, offline, and edge intelligence [24]. Devices become more powerful when they are able to make decisions based on the local data instead of waiting for every decision from the cloud. Thus, it should be ensured that the IoT platform has the capacity to support new topologies and utilize edge intelligence [35].

## IV. RESEARCH METHOD AND DATA COLLECTION

The data collection work began with the literature survey of the research articles and publications related to IoT and its platforms published in different journals, conferences, and books. About 200 articles were searched for the selected topic and 46 were selected for the study. The selected articles are searched in the IEEE, ACM, and Scopus databases, Google Scholar and some websites are searched and referenced. Twenty-one key factors of an IoT platform were identified from the literature. The preliminary analysis has been published by the authors in a conference paper [1]. Here we have selected five platforms as presented in Table I, to be discussed later. Relevant parts are summarized again here for clarity within the rest of this article. The collected key factors of an IoT platform were verified and categorized using the Delphi method, which is an interactive process to collect and distill the data using the judgments of experts using a feedback loop. This method is a flexible research technique that can be used to successfully explore new concepts inside and outside the information system body of knowledge [36].

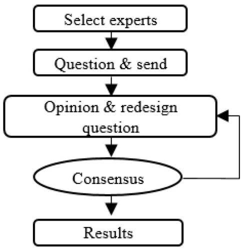



Fig. 2. Schematic of the Delphi method for verification and categorization.

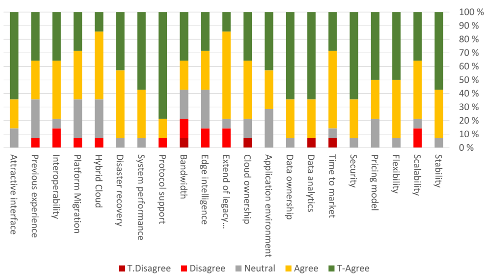



Fig. 3. Experts opinion in the first round.

We employed here a two-round Delphi study, as shown in Fig. 2, during the first round of the Delphi study 15 experts from three different universities were selected based on their experience in the IoT field. A questionnaire was designed based on 21 questions related to the key factors of IoT platforms as show in Table V in the Appendix. A five-point Likert rating scale was used: 1) totally disagree; 2) disagree; 3) neutral; 4) agree; and 5) totally agree. The questionnaire was sent to the experts by email to be answered within two weeks. Fourteen experts replied and the response percentage was 93%. The experts' opinion of the first round is shown in Fig. 3. In the first round, the agreed percentage is 80%, disagree percentage is 6%, and the neutral percentage is 14%.

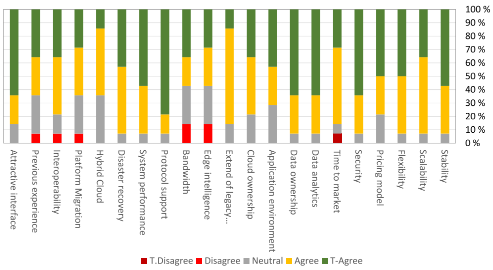



Fig. 4. Experts opinion in the second round.

There was little conflict between the opinions of the experts about the first round questions as shown in Fig. 3, the second-round questionnaire was designed based on the experts' opinion of the first round. The original questions were the same as before, only the summary of the experts' opinion of the first round was subsequently sent to the same experts. In the second round, 14 experts replied and the response percentage was 93%. In the second round, some of the experts have changed their opinion based on the summary of opinions of the first round. The result is shown in Fig. 4. The agreed percentage was then 81%, disagree percentage is 4%, and the neutral percentage is 15%. The results of both the rounds are shown in Table II. Note that for simplicity, we have merged "totally agree" and "agree" to "Agree," and "totally disagree" and "disagree" values to "Disagree."

The importance of all the 21 factors of an IoT platform was categorized into three categories in the light of experts' opinions. Factors with agree percentage up to 79% and above are considered very important, factors with agree percentage between 78% and 64% are considered somewhat important, and factors with agree percentage less than 60% are considered as less important. According to experts opinions, the factors stability, security, protocol support, system performance, disaster recovery, data analytics, scalability, flexibility, data ownership, extend of legacy architecture, pricing model, interoperability, attractive interface, cloud ownership, and time to market were considered as the most important factors. Four factors, application environment, hybrid cloud, platform migration, and previous experience were considered as somehow important, and two factors edge intelligence and bandwidth were considered as less important.

## V. Proposed Method to Compare IoT Platforms

### A. Identifying IoT Platforms

IoT applications need a platform to run smoothly and handle the data so that companies can take future decisions based on the data processed by the IoT platform [37]. Hundreds of IoT platforms are available and finding the most suitable IoT platform for a specific IoT application is becoming increasingly difficult. The problem is compounded by a lack of experience and knowledge, and in some cases, a company may select a platform without adequate requirements analysis, which later leads to problems [7]. When developing an IoT application for business needs, IoT platforms are the first place that can provide the facilities for deploying and running the business application [38]. There is high competition between various IoT platforms in the market. In this article, we have arbitrarily selected five well-known platforms based on information collected by specialized websites[1] and reports about their market share (e.g., [39]), which we will then use to demonstrate our methodology to support selecting appropriate platform solutions. The selected IoT platforms are: Amazon Web services (AWS) IoT, Google cloud IoT platform (GCP), Microsoft Azure IoT suite, IBM Watson IoT, and Oracle IoT. Their basic technical features are compared in terms of security, data analytics, protocols, visualization tool, data format, and application environment, as shown in Table I. Next, we will briefly present these platforms based on their own descriptions and other specialized references, trying to mimic how organizations collect information for selecting the service provider.

*Amazon Web Services (AWSs):* AWS was launched in 2006 and is the leading platform with 33% market share in 2018 [39]. AWS provides storage space, compute capability, data management, and other infrastructure resources [38]. It also offers artificial intelligence (AI) services [15]. AWS has customers, such as Dropbox, Netflix, and Philips [46].

*Microsoft Azure:* Azure was launched in 2010 and had in 2018 a market share of 24% [39]. It is capable of data gathering, processing, storing, and using analytics. It also allows IoT applications to work in a two-way communication [16]. Azure has customers, such as Apple-iCloud, EasyJet, and Xerox.

*Google Cloud IoT Platform:* GCP was launched in 2008 having market shares of 12% in 2018 [39]. GCP uses cloud and edge computing. It offers data analytics and machine learning while employing Google Maps to track the assets' positions. GCP has customers, such as PayPal and Bloomberg [18], [47].

*IBM Watson IoT Platform:* IBM Watson had a market share of 18% in 2018 [39] and provides connectivity, analysis, device management, and information management [19], [48]. IBM Watson employs two-way communication with the end user and also uses blockchain services. The main customers are STAPLES and AUTODESK.

*Oracle IoT Platform:* Oracle[2] offers acquisition, analysis, and integration of data [49], also using edge analytics [17]. The main customers of the Oracle IoT platform are Softbang LLC and Anson McCade.

### B. Proposed Method

Previous studies have also sought to identify relevant factors for selecting an IoT platform for business. Table III summarizes these. While there is a considerable overlap none of these prior studies have identified the number or granularity of factors as in our approach. Therefore, while they are no doubt suitable for the specific applications and domains within which they were developed, we have aimed to create a more general

[1]For example, https://internetofthingswiki.com/top-20-iot-platforms/634/.

[2]Its market share is not available in [39]; our selection was based on https://www.zdnet.com/article/top-cloud-providers-2018-how-aws-microsoft-google-ibm-oracle-alibaba-stack-up/.

TABLE II
RESULTS OF DELPHI STUDY BOTH ROUNDS. THE MEAN AND MEDIAN ARE TAKEN FROM THE AGREED VALUES

| | | **Survey round 1** | | | | | **Survey round 2** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **F** | **Factor** | Mean | Median | Disagree % | Neutral % | Agree % | Mean | Median | Disagree % | Neutral % | Agree % |
| F1 | Scalability | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 0% | 7% | 93% |
| F2 | Flexibility | 4 | 4.5 | 0% | 7% | 93% | 4 | 4.5 | 0% | 7% | 93% |
| F3 | Data anlytics | 4 | 5 | 7% | 0% | 93% | 5 | 5 | 0% | 7% | 93% |
| F4 | Disaster recovery | 4 | 4 | 0% | 7% | 93% | 4 | 4 | 0% | 7% | 93% |
| F5 | Stability | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F6 | Security | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F7 | Data ownership | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F8 | Protocol support | 5 | 5 | 7% | 0% | 93% | 5 | 5 | 0% | 7% | 93% |
| F9 | System performance | 5 | 5 | 0% | 7% | 93% | 5 | 5 | 0% | 7% | 93% |
| F10 | Time to market | 4 | 4 | 7% | 7% | 86% | 4 | 4 | 7% | 7% | 86% |
| F11 | Legacy architecture | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 0% | 14% | 86% |
| F12 | Attractive interface | 5 | 5 | 0% | 14% | 86% | 5 | 5 | 0% | 14% | 86% |
| F13 | Pricing model | 4 | 4.5 | 0% | 21% | 79% | 4 | 4.5 | 0% | 21% | 79% |
| F14 | Cloud ownership | 4 | 4 | 7% | 14% | 79% | 4 | 4 | 0% | 21% | 79% |
| F15 | Interoperability | 4 | 4 | 14% | 7% | 79% | 4 | 4 | 7% | 14% | 79% |
| F16 | App. environment | 4 | 4 | 0% | 29% | 71% | 4 | 4 | 0% | 29% | 71% |
| F17 | Hybrid cloud | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 0% | 36% | 64% |
| F18 | Platform migration | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 7% | 29% | 64% |
| F19 | Previous experience | 4 | 4 | 7% | 29% | 64% | 4 | 4 | 7% | 29% | 64% |
| F20 | Edge intelligence | 4 | 4 | 14% | 29% | 57% | 4 | 4 | 14% | 29% | 57% |
| F21 | Bandwidth | 4 | 4 | 21% | 21% | 57% | 4 | 4 | 14% | 29% | 57% |
| **–** | **Percentage** | – | – | **6%** | **14%** | **80%** | – | – | **4%** | **15%** | **81%** |

TABLE III
FACTORS FOR SELECTING IoT PLATFORM FACTORS

| Source | Year | Factors |
|---|---|---|
| [40] | 2018 | stability, time to market, pricing model, protocols, data analytics |
| [7] | 2018 | stability, scalability, flexibility, pricing model |
| [20] | 2015 | Security |
| [21] | 2015 | interoperability |
| [41] | 2018 | interoperability, security, reliability, protocols, data analytics, interactive interface. |
| [42] | 2016 | protocols, data analytics, scalability, security. |
| [43] | 2019 | interactive interface,interoperability,security,connectivity services,device management. |
| [44] | 2018 | protocol, security, bandwidth, application environment, cost. |
| [45] | 2019 | scalability, stability. |

TABLE IV
REFLECTING THE 21 KEY IoT PLATFORM FEATURES IN THE FIVE MAIN IoT PLATFORMS

| **Factors** | **AWS** | **Azure** | **Google cloud** | **IBM Watson** | **Oracle IoT** |
|---|---|---|---|---|---|
| Scalability | yes | yes | yes | yes | yes |
| Flexibility | yes | - | yes | - | yes |
| Data analytic | yes | yes | yes | yes | yes |
| Disaster recovery | yes | yes | no | no | no |
| Stability | yes | yes | yes | - | - |
| Security | high | high | high | high | high |
| Data ownership | - | yes | - | - | - |
| Protocol support | yes | yes | - | yes | yes |
| System performance | yes | - | yes | yes | - |
| Time to market | yes | yes | - | - | yes |
| legacy architecture | yes | - | - | - | yes |
| Attractive interface | yes | yes | - | no | - |
| Pricing model | bad | bad | good | - | - |
| Cloud ownership | yes | yes | yes | - | yes |
| Interoperability | yes | - | - | - | yes |
| App. environment | yes | yes | yes | yes | yes |
| Hybrid cloud | yes | yes | - | - | - |
| Platform migration | yes | yes | - | - | - |
| Previous experience | yes | yes | - | - | - |
| Edge intelligence | yes | yes | yes | - | yes |
| Bandwidth | - | - | good | - | - |

approach that can be more widely used across all these cases. To show how our general framework can be applied to assessing and choosing an IoT platform, in this article, we have selected the top five IoT platforms based on market share. We have compared these IoT platforms according to the 21 key IoT platform factors that we have identified from the literature

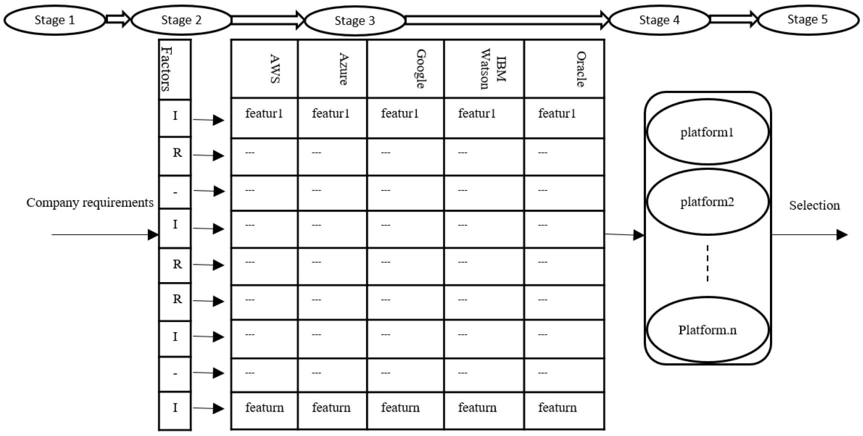



Fig. 5. Comparing key factors with the features offered by the IoT platform.

TABLE V
QUESTIONS USED IN SURVEY, DURING THE DELPHI METHOD

| Q# | Survey question |
|---|---|
| Q1 | What is your opinion about the importance of stability of IoT platform? |
| Q2 | What is your opinion about the importance of Scalability of the enterprise of IoT platform? |
| Q3 | Do you think that IoT platform should be flexible with the advancement of technologies? |
| Q4 | Do you think it is important to know about the pricing models before selecting IoT platform? |
| Q5 | Do you think IoT platform should provide security at both the ends, software and hardware? |
| Q6 | Do you think IoT platform can reduce Time to market for the business? |
| Q7 | Do you think IoT platform should support the basic descriptive, predictive and perspective analytics? |
| Q8 | Do you think it is important to know who will own the data collected by IoT platform? |
| Q9 | Is it important to know the application environment of IoT platform? |
| Q10 | Do you think it is important to know the Ownership of cloud infrastructure? |
| Q11 | Do you think extend of legacy architecture in IoT platform is important? |
| Q12 | Do you think Edge intelligence is important for IoT platform? |
| Q13 | Do you think IoT platform needs high bandwidth networking? |
| Q14 | Do you think it is important for IoT platform to support new Protocols and its updated versions? |
| Q15 | Do you think the IoT platform venders should implement some steps to keep System performance high? |
| Q16 | Do you think the IoT platform providers should have some dedicated infrastructure to handle customer data if there is some problem in IT infra? |
| Q17 | Do you think Hybrid cloud is important for IoT platforms? |
| Q18 | Do you think IoT platform providers should provide facilities to customers for any possible migration to other IoT platform in future? |
| Q19 | Do you think IoT platform Interoperability will enable the organization to get higher productivity? |
| Q20 | Is it necessary to check the previous experience of IoT platform, before selection? |
| Q21 | Is it necessary that user interface of the IoT Platform should be simple and attractive? |

and verified using the Delphi study. We have compared these 21 key factors with the features provided by those selected five IoT platforms as shown in Table IV.

More specifically, the entries of Table IV have the following meaning related to the specific feature to be considered: "yes" means the feature is available, "high" indicates strong, "bad" shows weak, "good" indicates that the feature is very good, "-" shows that the feature is unknown, and "no" indicates that the feature is not available in the platform. In order to identify and fill the features of the selected five IoT platforms, different articles [1], [7], [38], [40], [47], [49], [50] have been studied from many databases. Some websites [15]–[19], [23] have been used, especially the websites of those selected IoT platforms. A few white papers [51] have also been studied.

The framework for the selection of an IoT platform is illustrated in Fig. 5 as a schematic of the selection procedure. The whole process consists of five stages. In the first stage, the company finalizes their business requirements. In the second stage, the company requirements are applied to prioritizing, which factors are required (*R*), important (*I*), and not required (-) for this business context. In the third stage, the *R* and *I* factors are compared with the features provided by the five selected IoT platforms. The IoT platform/s that provide a maximum of the features as compared to the requirements are selected and shifted to stage four. In stage four, there might be one or many IoT platforms that match the required and important factors. Stage five is the decision, which is explained next.

If there is one IoT platform that provides the most required and important features, then the same IoT platform can be selected for the business application. But, if there are multiple IoT platforms providing these features, then the company may choose an IoT platform based on the comparison of their match to *I* factors, such as pricing, time to market, etc., and select a suitable IoT platform for their business needs. There might also be chances that none of the platforms provide all of the required features; this might indicate that new platforms should be selected and evaluated accordingly.

The five stages of the framework are explained in a simple example. A selling company is interested to be on the Web and uses IoT applications for its business. Initially they were interested to learn the components of IoT to understand what is IoT and how it works. Second, they were interested to know what is an IoT platform and what are the key factors of an IoT platform. They need an IoT platform for their business application but they do not know which platforms are providing what features and which one is the best. When the company has the knowledge of IoT, IoT platform factors and the features those platforms are providing then in stage 1, the company go through each of the 21 factors that have been identified as important in choosing a platform and use this to help them to formulate their business requirements. In stage 2, these factors are prioritized as being either required (*R*), important (*I*), or not required (-) for their business needs. They find that their required factors to consider (*R*) are scalability, time to market, and flexibility. Their important factors to consider (*I*) are pricing and interoperability. In stage 3, the *R* and *I* factors are compared with identified features of IoT platforms. AWS and Oracle are the platforms that are known to match all required features; both have the feature of interoperability, but AWS has the worse pricing model while one for Oracle is unknown. The company may request the pricing model from Oracle and then choose based on this.

## VI. Conclusion

The aim of this article is to build an objective methodology to support organizations to select the most suitable IoT platform based on their specific needs. To do so, we first reviewed the building blocks of IoT explaining how they are combined to perform specific tasks. Second, we identified 21 key factors of IoT platforms from the literature and then verified with the expert's opinion using Delphi studies. Finally, we have designed a theoretical framework for the selection of the IoT platform and tested it in five well-known examples. This article then provides a general framework to select the most suitable IoT platform for a specific organization by comparing its specific requirements with the features offered by the different platforms. As future work, we expect to evaluate the IoT platforms in different vertical cases, such as energy and Industry 4.0. Our goal is to build an automated procedure that also includes the possibility of weighing the factors based on interviews with experts, developers, and programmers from that particular domain.

## Appendix

Table V shows the questions used to carry out the Delphi method employed to validate the proposed 21 key factors.

## References


[1] M. Ullah and K. Smolander, "Highlighting the key factors of an IoT platform," in *Proc. 42nd Int. Convent. Inf. Commun. Technol. Electron. Microelectron. (MIPRO)*, 2019, pp. 901–906.

[2] C. Greer, M. Burns, D. Wollman, and E. Griffor, *Cyber-Physical Systems and Internet of Things*, NIST SP 1900-202, NIST, Gaithersburg, MA, USA, 2019.

[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[4] X. Jia, Q. Feng, T. Fan, and Q. Lei, "RFID technology and its applications in Internet of Things (IoT)," in *Proc. 2nd Int. Conf. Consum. Electron. Commun. Netw. (CECNet)*, 2012, pp. 1282–1285.

[5] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, "A gap analysis of Internet-of-Things platforms," *Comput. Commun.*, vols. 89–90, pp. 5–16, Sep. 2016, doi: 10.1016/j.comcom.2016.03.015.

[6] Zana Diaz Williams. (2017). *IoT Platform Company List 2017*. [Online]. Available: https://iot-analytics.com/iot-platforms-company-list-2017-update/

[7] H. Hejazi, H. Rajab, T. Cinkler, and L. Lengyel, "Survey of platforms for massive IoT," in *Proc. IEEE Int. Conf. Future IoT Technol. Future IoT*, Jan. 2018, pp. 1–8.

[8] J. Guth, U. Breitenbucher, M. Falkenthal, F. Leymann, and L. Reinfurt, "Comparison of IoT platform architectures: A field study based on a reference architecture," in *Proc. Cloudification Internet Things (CIoT)*, 2017, pp. 1–6.

[9] C. Okoli and K. Schabram, "A guide to conducting a systematic literature review of information systems research," in *Proc. SSRN*, vol. 10, 2011, p. 2010.

[10] N. Koshizuka and K. Sakamura, "Ubiquitous ID: Standards for ubiquitous computing and the Internet of Things," *IEEE Pervasive Comput.*, vol. 9, no. 4, pp. 98–101, Oct.–Dec. 2010.

[11] Y. Song. (2013). *Security in Internet of Things*. [Online]. Available: http://kth.diva-portal.org/smash/get/diva2:702223/FULLTEXT01.pdf%0Ahttp://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-142451

[12] S. Mandal. (2016). *Internet of Things (IoT)—Part 2 (Building Blocks & Architecture)*. [Online]. Available: https://www.c-sharpcorner.com/UploadFile/f88748/internet-of-things-part-2/

[13] M. Gigli and S. Koo, "Internet of Things: Services and applications categorization," *Adv. Internet Things*, vol. 1, no. 2, pp. 27–31, 2011. [Online]. Available: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ait.2011.12004

[14] F. H. Mohammed and R. Esmail, "Survey on IoT services: Classifications and applications," *Int. J. Sci. Res.*, vol. 4, no. 1, pp. 2124–2127, 2015.

[15] Amazon Web Services. (2019). *IoT Applications & Solutions—What Is the Internet of Things (IoT)?—AWS*. [Online]. Available: https://aws.amazon.com/iot/

[16] Microsoft Azure IoT. (2019). *Microsoft Azure*. [Online]. Available: https://azure.microsoft.com/en-in/pricing/details/iot-hub/

[17] Internet of Things. (2019). *Internet of Things (IoT)—Oracle*. [Online]. Available: https://www.oracle.com/internet-of-things/

[18] Google Cloud. (2019). *Google Cloud IoT—Fully Managed IoT Services—Google Cloud*. [Online]. Available: https://cloud.google.com/solutions/iot/

[19] IBM Watson IoT Platform. (2019). *IBM Watson IoT Platform—Overview—Finland*. [Online]. Available: https://www.ibm.com/fi-en/marketplace/internet-of-things-cloud

[20] S. Narula, A. Jain, and Prachi, "Cloud computing security: Amazon Web service," in *Proc. 5th Int. Conf. Adv. Comput. Commun. Technol.*, 2015, pp. 501–505. [Online]. Available: http://ieeexplore.ieee.org/document/7079135/

[21] V. Gazis *et al.*, "A survey of technologies for the Internet of Things," in *Proc. 11th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, 2015, pp. 1090–1095.

[22] J. Lee. (2018). *How to Choose the Right IoT Platform: The Ultimate Checklist | Hacker Noon*. Accessed: Sep. 20. 2020. [Online]. Available: https://hackernoon.com/how-to-choose-the-right-iot-platform-the-ultimate-checklist-47b5575d4e20

[23] IOTIFY. *Top 10 Criteria to Choose the Best IoT Cloud Platform*. Accessed: Oct. 26, 2018. [Online]. Available: https://iotify.io/top-10-selection-criteria-for-your-iot-cloud-platform/

[24] Jeffrey Lee. (2018). *The 6 Complexities of Building a Managed IoT Platform*. [Online]. Available: https://hackernoon.com/the-6-complexities-of-hosting-a-managed-iot-service-b7696eea52ba

[25] Igor Ilunin. (2017). *How to Choose Your IoT Platform—Should You Go Open-Source?*. [Online]. Available: https://medium.com/iotforall/how-to-choose-your-iot-platform-should-you-go-open-source-23148a0809f3

[26] Calum McClelland. *Leverege—How to Choose the Best IoT Platform*. Accessed: Sep. 19, 2018. [Online]. Available: https://www.leverege.com/blogpost/how-to-choose-the-best-iot-platform

[27] J. Khakurel. *Enhancing the Adoption of Quantified Self-Tracking Devices*. Accessed: Dec. 29, 2018. [Online]. Available: http://urn.fi/URN:ISBN:978-952-335-319-0

[28] E. Lamarre and B. May. (2017). *Making Sense of Internet of Things Platforms—McKinsey*. [Online]. Available: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/making-sense-of-internet-of-things-platforms

[29] J. P. Matthew. (2016). *Evaluating and Choosing an IoT Platform—O'Reilly Media*. [Online]. Available: https://www.oreilly.com/learning/evaluating-choosing-iot-platform

[30] T. Salman. (2015). *Internet of Things Protocols and Standards*. [Online]. Available: https://www.cse.wustl.edu/ jain/cse570–15/ftp/iot_prot/

[31] K. Vandikas and V. Tsiatsis, "Performance evaluation of an IoT platform," in *Proc. 8th Int. Conf. Next Gener. Mobile Appl. Services Technol. (NGMAST)*, 2014, pp. 141–146.

[32] J. Koo and Y. Kim, "Interoperability of device identification in heterogeneous IoT platforms," in *Proc. 13th Int. Comput. Eng. Conf. (ICENCO)*, Cairo, Egypt, 2017, pp. 26–29, doi: 10.1109/ICENCO.2017.8289757.

[33] G. Xiao, J. Guo, L. D. Xu, and Z. Gong, "User interoperability with heterogeneous IoT devices through transformation," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1486–1496, May 2014.

[34] M. Al-Zihad, S. A. Akash, T. Adhikary, and M. A. Razzaque, "Bandwidth allocation and computation offloading for service specific IoT edge devices," in *Proc. 5th IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC)*, 2018, pp. 516–519.

[35] D. Loghin, L. Ramapantulu, and Y. M. Teo, "On understanding time, energy and cost performance of Wimpy heterogeneous systems for edge computing," in *Proc. IEEE 1st Int. Conf. Edge Comput. (EDGE)*, 2017, pp. 1–8.

[36] G. J. Skulmoski, F. T. Hartman, and J. Krahn, "The Delphi method for graduate research," *J. Inf. Technol. Educ. Res.*, vol. 6, pp. 1–21, Jan. 2007.

[37] D. Elizalde. (2017). *Should I Use an IoT Platform?—Reduce Cost and Boost Quality—IoT for All*. [Online]. Available: https://www.iotforall.com/should-i-use-an-iot-platform/

[38] B. Nakhuva and T. Champaneria, "Study of various Internet of Things platforms," *Int. J. Comput. Sci. Eng. Survey*, vol. 6, no. 6, pp. 61–74, 2015.

[39] R. Nv. (2018). *Cloud Growth Rate Increased Again in Q1; Amazon Maintains Market Share Dominance—Synergy Research Group*. [Online]. Available: https://www.srgresearch.com/articles/cloud-growth-rate-increased-again-q1-amazon-maintains-market-share-dominance

[40] P. Agarwal and M. Alam. (2018). *Investigating IoT Middleware Platforms for Smart Application Development*. [Online]. Available: http://arxiv.org/abs/1810.12292

[41] Y. Kondratenko, G. Kondratenko, and I. Sidenko, "Multi-criteria decision making for selecting a rational IoT platform," in *Proc. IEEE 9th Int. Conf. Depend. Syst. Services Technol. (DESSERT)*, 2018, pp. 147–152.

[42] P. Ganguly, "Selecting the right IoT cloud platform," in *Proc. Int. Conf. Internet Things Appl. (IOTA)*, 2016, pp. 316–320.

[43] P. Victor. (2019). *IoT Platform Selection Guide—DZone IoT*. [Online]. Available: https://dzone.com/articles/iot-platform-selection-guide

[44] Maker.io. (2018). *How Do You Pick the Right IoT Platform?*. [Online]. Available: https://www.digikey.fi/en/maker/blogs/2018/how-do-you-pick-the-right-iot-platform

[45] A. A. Ismail, H. S. Hamza, and A. M. Kotb, "Performance evaluation of open source IoT platforms," in *Proc. IEEE Global Conf. Internet Things (GCIoT)*, 2019, pp. 1–5.

[46] A. Rojas. (2017). *A Brief History of AWS—The Media Temple Blog*. [Online]. Available: https://mediatemple.net/blog/news/brief-history-aws/

[47] S. Challita, F. Zalila, C. Gourdin, and P. Merle, "A precise model for Google cloud platform," in *Proc. IEEE Int. Conf. Cloud Eng. (IC2E)*, 2018, pp. 177–183.

[48] A. Salami and A. Yari, "A framework for comparing quantitative and qualitative criteria of IoT platforms," in *Proc. 4th Int. Conf. Web Res. (ICWR)*, 2018, pp. 34–39.

[49] P. P. Ray, "A survey of IoT cloud platforms," *Future Comput. Informat. J.*, vol. 1, nos. 1–2, pp. 35–46, 2017, doi: 10.1016/j.fcij.2017.02.001.

[50] M. Ammar and B. Crispo, "slimIoT: Scalable lightweight attestation protocol for the Internet of Things," in *Proc. DSC*, 2018, pp. 1–8.

[51] "Google cloud security white papers," Google, Mountain View, CA, USA, Mar. 2018, pp. 1–97.

**Mehar Ullah** (Member, IEEE) received the B.S. degree in information technology from Iqra National University, Peshawar, Pakistan, in 2004, and the master's degree in software engineering from the Lappeenranta–Lahti University of Technology, Lappeenranta, Finland, in 2014, where he is currently pursuing the Doctoral degree.

His main research field is IoT and cyber-physical systems specially for industrial applications.

**Pedro H. J. Nardelli** (Senior Member, IEEE) received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Campinas, Brazil, in 2006 and 2008, respectively, and the Doctoral degree from the University of Oulu, Oulu, Finland, and the State University of Campinas following a dual degree agreement in 2013.

He is currently an Assistant Professor (tenure track) of Internet of Things (IoT) in energy systems with LUT University, Lappeenranta, Finland, and holds a position of Academy of Finland Research Fellow with a project called Building the Energy Internet as a large-scale IoT-based cyber–physical system that manages the energy inventory of distribution grids as discretized packets via machine-type communications. He leads the Cyber-Physical Systems Group with LUT and is a Project Coordinator of the CHIST-ERA European Consortium Framework for the Identification of Rare Events via Machine Learning and IoT Networks. He is also an Adjunct Professor with the University of Oulu in the topic of "Communications Strategies and Information Processing in Energy Systems." His research focuses on wireless communications particularly applied in industrial automation and energy systems.

Dr. Nardelli received the Best Paper Award of IEEE PES Innovative Smart Grid Technologies Latin America 2019 in the track "Big Data and Internet of Things." More details can be found in https://sites.google.com/view/nardelli/.

**Annika Wolff** received the B.Sc. degree in cognitive science from the University of Hertfordshire, Hatfield, England, in 2000, and the Ph.D. degree in knowledge modeling and HCI from Open University, Milton Keynes, England, in 2016.

She is a Postdoctoral Researcher with LUT University, Lappeenranta, Finland, and a Visiting Research Fellow with Open University, Milton Keynes, U.K. Her research is in the newly emerging field of human–data interaction, at the intersection between complex data, machine, and human learning. Her research focuses on how people make sense of, interact with and design from complex data, such as derived from IoT and within smart cities.

Dr. Wolff is currently the Co-I of the H2020 (SWAFS) Project ParCos on the Participatory Communication of Science, and the Conference Planning Chair for the European Society for Socially Embedded Technologies as well as an Associate Editor of the *International Journal of Human–Computer Studies*.

**Kari Smolander** received the Licentiate degree from the University of Jyväskylä, Jyväskylä, Finland, in 1993, and the Ph.D. degree in computer science from LUT University, Lappeenranta, Finland, in 2003.

He is a Professor of software engineering with the School of Engineering Science, LUT University and an Adjunct Professor with the Department of Computer Science, Aalto University, Espoo, Finland. He is the Head of the Software Engineering Department, LUT University. He has published more than 180 refereed research papers in international journals and conferences. His current research interests include change in software and systems development practices and organizations.

# Publication II

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P.

**Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data**

Reprinted with permission from
*IEEE Access*
2022, doi: 10.1109/ACCESS.2022.3215167.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2017.DOI*

# Industrial Energy Management System: Design of a Conceptual Framework using IoT and Big Data

**MEHAR ULLAH[1], ARUN NARAYANAN[1], ANNIKA WOLFF[2], AND PEDRO NARDELLI [1], (Senior Member, IEEE)**

[1]Department of Electrical Engineering, School of Energy Systems, LUT University, Lappeenranta, Finland
[2]Department of Software Engineering, School of Engineering Science, LUT University, Lappeenranta, Finland

Corresponding author: Mehar Ullah (e-mail: mehar.ullah@lut.fi)

This paper is partly supported by Academy of Finland via (a) FIREMAN consortium CHIST-ERA/n.326270, and (b) EnergyNet Research Fellowship n.321265/n.328869

**ABSTRACT** Industrial activities consume a large portion of the total energy demand worldwide and thus significantly contribute to greenhouse gas emissions. One of the most effective ways to reduce energy consumption in the industrial sector is to implement an energy management system. Current research into Industrial Energy Management System (IEnMS) remains insufficient, and to the best of our knowledge, a holistic framework for an IEnMS using the Internet of Things (IoT) and big data does not exist. This paper provides a comprehensive systematic literature review of the existing academic publications on IEnMS from where the main requirements and components of an IEnMS are identified. We further verify this study by conducting a detailed survey with specialized employees of ten (10) large companies to acquire expert opinion about using the modern technologies like IoT, big data, and data analytics in IEnMS. We have then proposed a theoretical framework for the IEnMS using IoT, big data and data analytics to construct an effective cyber-physical system architecture including steps from data acquisition to the end-user decision-making process. These findings demonstrate how the suggested framework provides an objective methodology for selecting the most appropriate IEnMS for various businesses based on their specific needs.

**INDEX TERMS** **Energy management systems, big data, IoT**

## I. INTRODUCTION

CLIMATE change is a major global issue that poses a threat to humanity's health and safety. It is critical to not only incorporate renewable energy sources but also the latest technologies to boost energy efficiency in order to positively contribute to combat such an urgent issue, and thus, help to protect current and future generations [1]. There is a strong vision that energy sector needs a significant shift from a fossil-fuel-dominated system towards one dominated by more environmental friendly sources.

One approach is to modify use patterns and increase energy efficiencies in electrified sectors such as residential, commercial and industrial [2]. Remarkably, industries do consume a tremendous amount of energy, estimated as 42.3% percent of all the energy produced globally [3]. It then becomes critical to develop and implement energy efficiency and management policies tailored to the industrial sector's specific issues.

In practical terms, industries should be encouraged to prioritize the management and operation of their own energy systems to ensure their long-term sustainability. One approach is to explicitly deploy energy management and control policies as a viable strategy for reducing energy consumption, associated energy expenditures, and carbon emissions in the workplace. Because energy prices have not been traditionally acted as a strong constraint to the production size in large scale corporations, and the decrease of greenhouse gas emissions were not an aim in the 20$^{\text{th}}$ century, industries have usually grown without any specific focus on energy-related aspects, including energy efficiency in general and environmental sustainability related to energy sources in particular. However in the 21$^{\text{st}}$ century, the situation has changed: energy prices indicates a strong tendency to increase together with an unstable geopolitical situations in different producing countries, regulatory issues with nuclear sources and a strong push towards emissions decrease coming from the poor environmental footprint of fossil fuel energy carriers. As a result, the industrial sector has begun to place greater emphasis on energy management targeting at deploying energy efficient solutions by decreasing of consumption, or by using intelligent production schedule. This may come from the deployment of more efficient equipment, or from improvements in the process design. These aspects are also combined with the use of new distributed energy sources like solar and wind, which can be combined with new industrial processes related to Power-to-X as long term storage units. In this context of economical and political uncertainties, many corporation are today developing long-term energy strategies to meet the specified energy usage targets and values.

With the fast pace development of dedicated Information and Communication Technologies (ICTs), the widespread of Internet of Things (IoT) is potentially beneficial in this domain because of the possibility to logically interconnect production elements through advanced sensors, mainly after the push towards digitalization that took place during the pandemics. IoT can play a major role in energy efficient utilization physical production systems by serving as the key enabler for an Industrial Energy Management System (IEnMS) that includes improved situational awareness in monitoring, increased intelligence and automation in operation and control, and increased efficiency in energy dispatch and management to enable cheaper and cleaner energy usage.

In particular, the IEnMS can provide via specific sensors and/or actuators different ways of intervene in the energy demand of industries. For example, a given industrial process can be automatically scheduled to operate during the time that energy from solar PVs are available. Other case is when the IEnMS monitors the power demanded by different industrial tools, and thus, it is capable of indentifying the ones that demand more energy; this might guide the responsible personnel to replace such tools to most energy efficient ones. A last example is related to the flexibility of industrial loads that might be used by controllers and actuators that may follow operational signals to turn on or turn off (or decrease or increase the power demanded if this is a feature of the device/process). It is also important to highlight that home energy management systems (HEMS), although may share some of the basic aspects of the IEnMS, has a much more simplified network architecture based on already existing local area networks, or already deployed cloud solutions, which also has less strict constrains if compared to industrial processes. Also, the way electricity markets are structured usually leads to a differentiation of industrial and residential consumers, which has different impact in the functioning of IEnMS and HEMS.

The complete process of regulating energy consumption, reducing greenhouse gas emissions, decreasing energy usage, and achieving considerable cost savings is implemented using an IEnMS. In this scenario, IEnMS can—and should—be implemented considering larger time horizons and different factors, which many times contradicts the usual operational approaches that aim at short-term minimization of costs or maximization of profits, only.

This contribution proposes a high-level architecture for IEnMS that incorporates IoT, big data processing and data analytics. Although this topic has been extensively studied [4]–[6] as the basis of HEMS, the focus on IEnMS implementations are still lacking. HEMS and IEnMS, in particular, have roughly similar goals, although they face quite different challenges because of their different concrete aims with clear diverse operational scales and needs. Our goal is to give a basic framework that different industrial players can use to choose the best appropriate IEnMS for their specific needs. To put it another way, this article expects to assist industries in doing a complete analysis of their own energy requirements and comprehending the major components of their energy management schemes in order to identify the greatest fit for their IEnMS.

In this context, the aim of this paper is to establish an energy-centric approach to organize industrial environments employing the state-of-the-art in ICTs. As argued above, the existing literature gives far too much attention to HEMS, leaving IEnMS in the second plane. Here, our aim is to fill part of this research gap by identifying general guidelines to construct and operate information systems for different industries with energy as the main focus. Specifically, this paper focuses on finding the answers to the following four questions:

1) What are the current energy management methods in the industrial sector?
2) What are the main requirements and components of Industrial Energy Management Systems?
3) What are the views of the energy employees in industries about the role of using the latest technologies like IoT and big data in Industrial Energy Management Systems?
4) How can a high-level architecture for a Industrial Energy Management System be designed that incorporates latest advanced technologies like IoT and big data?

To answer these questions, a detailed literature review is provided to identify the best practices in energy management in the industrial sector. The results of a survey with experts opinions whose objective was to identify the importance of the most advanced ICTs to support the energy management activities of their corporations is presented, revealing the willingness to use IEnMS as a tool to support decision-making in different aspects that concern energy consumption and efficiency, as well as environmental sustainability. Finally, from the literature review and the survey results, we have proposed a general framework to build an IEnMS considering IoT connectivity and big data analysis. Note, however, that our paper focuses on a high-level design, and thus, the specific details of which type of IoT devices and their respective communication protocols and data frames are out of our scope. We acknowledge that those aspects are essential in the deployment of any particular IEnMS, but our aim here is to provide general designing principles that should be in place so that the IEnMS would function properly.

The rest of the paper is organized as follows. Section II describes the background of the paper, related literature, and the identified research gap. Section III explains the Industrial Energy Management and how energy can be managed in industries. Section IV elucidates a more systematic approach toward industrial energy management: here, we define and explain an Industrial Energy Management System and explain in detail why a systematic approach is both necessary and beneficial for industrial energy management. Section V explains working of IoT and big data and contains the benefits of the combination of IoT and big data. Section VI presents the survey questions, results and discussion. In Section VII, we present our proposed theoretical framework for the Industrial Energy Management System using IoT and big data. Finally, in Section VIII, we discuss and conclude the paper.

## II. BACKGROUND

Numerous research on industrial decision-making processes related to energy-related concerns exist, with universal agreement on answers and techniques but considerable disagreements on the long-term repercussions of extreme change. Some companies believe that investing in Energy Management (EnM) programs will have a significant financial benefit, and that the financial impact will be the key factor motivating EnM program implementation [7], [8]. Other authors [9], [10] consider that energy policy, pricing, expertise, and mindset are all energy-related choice factors that influence EnM programs. More recently, some authors [1], [11]–[13] have shown that energy-related decisions are based on the strategic links between the organization's main business and goal with any investment. In today's industrial sector, the importance and necessity of implementing EnM for optimal energy use is a well-established reality, and most organizations acknowledge the need for a correctly employed energy management system [14]; As a result, many companies are enacting or have enacted some form of EnM policy. Nonetheless, there is still a misconception about the relationship between EnM and Energy Management System (EnMS), necessitating a thorough research program to further clarify the differences [15].

According to [16], The procedures and processes by which the organization strategically handles energy challenges and management are referred to as EnM. EnMS is the instrument that is utilized to put those practices and processes into action. Thus,for example EnMS must be applied in order to successfully build and apply EnM methods and processes aimed at reducing energy consumption, cost, and greenhouse gas emissions. Another study found that participation of top management as well as almost all of a company's employees in all energy-related activities is essential not just for a successful EnMS deployment, but also to ensure that the best EnM practices are implemented company-wide [17].

In the literature, EnM has a systematic approach that comprises of practices and activities that are considered supporting functions for an industrial EnM.Despite the fact that EnM is a critical goal at the moment, it has yet to be fully implemented in the industrial sector. This is due to the interdisciplinary character of EnM in industries, which includes numerous aspects other than economics and technology, such as social acceptance, political viewpoints, and managerial established procedures. To ensure the effective implementation of EnM, a type of EnMS applied to industries—Industrial EnMS (IEnMS)—is desirable, so that an objective means to monitor, plan, and regulate energy consumption and efficiency can be built, and the intended energy positive goals may be met. This is the phase that is frequently misunderstood, and as a result, the majority of the industrial sector fails to perform it effectively and correctly. Furthermore, while IoT and big data are essential enabling technologies that can be utilized to support complicated energy management processes, their precise roles in the literature have yet to be determined. These are the research gaps that have been identified in this study, and they are the focus of this paper. In the following sections, we'll first go through EnM in the context of industry, i.e., Industrial Energy Management (IEnM), and then show how an IEnMS can be utilized to accomplish efficient IEnM.

## III. INDUSTRIAL ENERGY MANAGEMENT

Growth in the industrial sector, as well as its consequent increase in energy consumption, is one measure of economic progress [18]. It is then critical to establish IEnM approaches that support the efficient use of energy while reducing carbon emissions to counteract the negative effects of such expansion [3]. In order to enhance their operational efficiency in this domain and thereby reduce their respective consumption with their related carbon emissions, many companies have turned their attention to energy-related issues in the more recent years. Our studies have revealed that there are few identifiable key components that IEnM should have: planning/strategy, operation/implementation, controlling, organization, and culture. as depicted in Fig. 1 [18]–[22]. The details of each of these components are presented next.

**Planning/Strategy**: The Planning/Strategy component is the initial phase of IEnM, and it is divided into three sections: (1) Formulation of a company's long-term energy policy that is conducive to a successful IEnM [23]. (2) Energy planning and goal-setting, in which a company establishes plans and sets goals for future energy usage [24], [25]. (3) Strategic energy risk management, in which firms assess any sort of energy-related risk and provide risk management strategies depending on the company's financial goals and risk tolerance [26].

**Operation/Implementation**: IEnM's second component is made up of three parts: (1) Energy efficiency measures

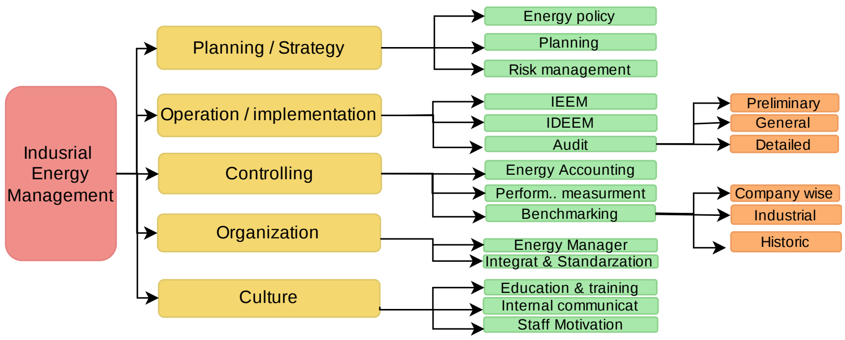



**FIGURE 1.** Components of Industrial Energy Management.

are implemented, in which businesses implement specialized energy projects and energy efficiency technologies in order to reduce electricity consumption [27]. (2) Investment decisions based on energy efficiency, in which corporations perform systematic economic calculations to determine the return on investment. [28]. (3) Energy audit, in which organizations involved in energy management activities assess the status quo on a regular basis and identify energy-saving opportunities.There are three sorts of audits:Preliminary audits, general audits, and detailed audits [27].

**Controlling**: The third component of IEnM, controlling, is divided into three parts: energy accounting, performance assessment, and benchmarking.Energy accounting is the practice of continuously analyzing and reporting energy consumption as well as monitoring energy efficiency [29]. IEnM includes performance measurement, which defines the key performance indicators (KPIs) for energy efficiency, which characterize the relationship between an activity and the amount of energy consumed [20]. Energy benchmarking is a performance-oriented activity that can be defined as the process for comparing energy efficiency between or within entities. Benchmarking is a useful tool for lowering energy consumption, prices, and emissions. There are three forms of benchmarking: company-level benchmarking, industrial benchmarking, and historical benchmarking [30].

**Organization**: There are two parts to the organization: (1) Appointment of an experienced energy manager to keep top management informed about the energy management's actions and progress; the energy manager and top management should have a good working relationship [31], [32]. (2) Integration and standardization, in which industrial organizations' energy management should be connected with their production management processes through the use of Information and Communication Technology (ICT) tools and standardization [29].

**Culture**: Culture is the fifth and final component of IEnM, and it is divided into two parts: (1) Training and education:Personnel with adequate basic education to meet the energy-usage standards are required by the energy manager, or training may be required [33]. (2) Employee motivation: Businesses must encourage employees to take an active role in improving energy efficiency, and they should frequently award technical and operational personnel [29], [34].

To achieve effective and efficient IEnM that incorporates the above mentioned components, industries have to implement a comprehensive, company-wide, energy management system. We argue here that this can be successfully performed by an Industrial Energy Management System (IEnMS). The key features that the IEnMS should have is presented in the next section below.

## IV. INDUSTRIAL ENERGY MANAGEMENT SYSTEM

An ICT solution developed to facilitate the deployment of an effective IEnM is known as an Industrial Energy Management System (IEnMS). An IEnMS monitors, analyzes, and manages the energy demand (and potentially its generation and storage) in a manufacturing plant following the schematic presented in Fig. 2. It is also utilized to diagnose issues like overuse and leaks throughout the entire facility. IEnMS helps then to dynamically modulate the (supply-demand) energy profiles based on the current state of the energy system and to improve the efficiency of the energy consumption as whole by indicating potential sources of faults and unexpected performance in terms of energy.

An IEnMS is designed for large scale industrial energy consumers to manage their energy usage considering that their specific type of loads are different than residential ones; in industrial environments the load flexibility and consumption levels are quite specific, considering that some industrial processes cannot be turned on and off like a household heating system. To develop an efficient IEnMS based on their own special needs, industries must follow a set of basic processes, which can be divided in the following phases: (i) defining an energy policy and assigning duties, (ii) emphasizing major energy consumers, (iii) setting measurable goals and targets, (iv) putting in place actions to accomplish the goals, (v) determining whether the actions are successful, and (vi) conducting continuous system evaluations.

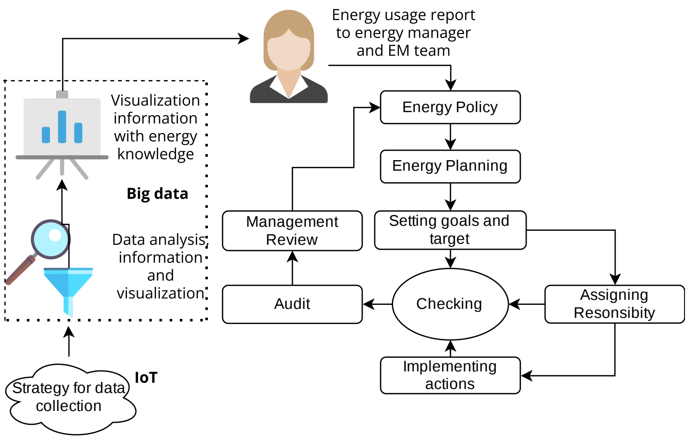



FIGURE 2. Parts of an Industrial Energy Management System

The IEnMS operation can be conceptualized as a cyber-physical system which required four basic steps, from data collection to semantic interpretation. In the following, these steps are systematically presented.

**(1) Build data collection strategy:** Real-time, precise, and granular data, as well as information on where and when the energy was utilized, and by which device, is collected (machine). Sensors/smart meters, sub-meters, and major energy consumer devices such as HVAC (heating, ventilation, and air conditioning) equipment, production lines, boilers, and other large energy consumer devices are used to collect data. The goal of this section is to keep track of the real-time data collection and figure out where the majority of the energy is being consumed.

**(2) Transform raw energy data into useful information:** The collected data is processed, evaluated, and turned into meaningful information during this step. Big data software is used to easily extract raw data from IoT devices and convert it into usable information in the form of user-friendly graphics. The raw data obtained might be linked to production levels, weather data, and other variables that influence the amount of energy consumed to generate the company's KPIs.

**(3) Assign responsibility, analyze data:** The information provided must be transformed into useful and meaningful reports during this phase; this can only be done by adding the information to the knowledge of facility, which can be done by an energy manager.The energy manager's job is to understand the information provided by the Energy Management System, combine it with the company's processes, and create appropriate targets.

**(4) Interpret the results, and agree to an action plan:**The energy manager gets access to the energy usage reports throughout this phase. The energy manager and his staff begin to speak with the departments in order to establish an energy strategy and a plan of action.

From the points just mentioned, the IEnMS may enable different benefits to industries; the main ones are listed next.

- **Reducing cost and saving energy:** IEnMS enables a continuous process that can lead to increased energy efficiency and productivity to lower energy expenses over time.
- **Planning future targets:** IEnMS supports the visualization of trends of energy demand, guide the definition of new strategic and operational goals.
- **Reducing greenhouse emissions:** IEnMS can be used to monitor greenhouse emissions from the plant operation (considering also the energy sources and emissions produced as byproduct of the industrial process itself). This is necessary for defining concrete targets related to emissions.
- **Increasing in machine lifespan:** Data collected from IEnMS can also be used for predictive maintenance, which can result in a longer life cycle for the monitored machines, also saving natural resources and decreasing investment costs.
- **Mitigation of risk related to fossil fuel:** IEnMS may help to integrate local energy sources by having local generation from solar and also power-to-X solutions developed for long-term storage like power-to-gas (PtG).
- **Improvement in company projects:** Data from IEnMS, can also be used in improved post-investment project performance, also indicating how energy use and carbon emissions tried to be minimized in a systematic way.

## V. WORKING OF IOT AND BIG DATA

IoT and big data refer to two different ICTs that have emerged in recent years; they have been developed in relative autonomy, but they are clearly interconnected [35]. IoT is playing a promising role in connecting the devices/machine

TABLE 1. Survey questions and the opinions Percentages

| S.No | Topic | Total Disagree | Disagree | Neutral | Agree | Totally Agree | Mean | Median |
|---|---|---|---|---|---|---|---|---|
| Q1 | An Industrial Energy Management System is good for energy saving energy saving | 0% | 0% | 0% | 30% | 70% | 4.7 | 5 |
| Q2 | An Industrial Energy Management System can reduce greenhouse gas emissions reduce greenhouse gas emissions | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q3 | There is lack of management awareness in industries for energy management for energy management | 0% | 20% | 30% | 30% | 20% | 3.5 | 3.5 |
| Q4 | An Energy Management System will provide energy saving opportunities saving opportunities | 0% | 0% | % | 50% | 50% | 4.5 | 4.5 |
| Q5 | The company's energy policy plays an important role in designing energy management system in designing energy management system | 0% | 10% | 10% | 20% | 60% | 4.3 | 5 |
| Q6 | Long term energy planning is important for industries | 0% | 0% | 0% | 30% | 70% | 4.7 | 5 |
| Q7 | An energy manager should create an energy management team management team | 0% | 0% | 30% | 40% | 30% | 4 | 4 |
| Q8 | It is important for the energy manager to be environment friendly environment friendly | 11% | 0% | 11% | 11% | 67% | 4.2 | 5 |
| Q9 | It is important that companies should display department-wise energy usage status using screens to motivate staff members wiseenergy usage status using screens to motivate staff members to save energy | 0% | 0% | 20% | 40% | 40% | 4.2 | 4 |
| Q10 | It is important for companies to give incentives and rewards to staff to encourage them to achieve energy targets tostaff to encourage them to achieve energy targets | 0% | 10% | 20% | 20% | 50% | 4.1 | 4.5 |
| Q11 | Companies need some renovation in the existing infrastructure to improve energy management infrastructure to improve energy management | 0% | 0% | 22% | 22% | 56% | 4.3 | 5 |
| Q12 | It is important for companies to add green energy like solar and wind energy to their existing energy usage solar and wind energy to their existing energy usage | 0% | 0% | 10% | 20% | 70% | 4.6 | 5 |
| Q13 | Companies should have a strong policy to reduce greenhouse gas emissions greenhouse gas emissions | 0% | 0% | 10% | 0% | 90% | 4.8 | 5 |
| Q14 | Installing sensors on machines, so that these machines can use IoT-based techniques to share real data with each other, will lead to improved energy efficiency and performance | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q15 | Using IoT and Big data in Industrial Energy Management Systems will facilitate timely identification and prevention of faults | 0% | 0% | 0% | 50% | 50% | 4.5 | 4.5 |
| Q16 | Companies should invest more in Industrial Energy Management Systems (IEnMS) | 0% | 0% | 10% | 60% | 30% | 4.2 | 4 |
| Q17 | IoT is helping to improve HVAC (heating, ventilation, and air conditioning) systems in manufacturing plants. | 0% | 0% | 20% | 50% | 30% | 4.1 | 4 |
| Q18 | IoT devices are capable of collecting a huge amount of real time data about different machines. Therefore, collecting and using Big Data is a good option for companies to perform real time data about different machines. | 0% | 0% | 11% | 33% | 56% | 4.4 | 5 |
| Q19 | IoT and Big Data will make data analysis and processing easier and will give energy information very quickly and this can be useful for business decisions in the future. | 0% | 0% | 11% | 22% | 67% | 4.6 | 5 |
| Q20 | Industries should use the latest IoT-enabled technologies in their Energy Management System to improve energy activities in their Energy Management System to improve energy activities like efficiency, performance, usage, cost etc | 0% | 0% | 10% | 20% | 70% | 4.6 | 5 |

together using data communication network so that they can share information with each other. These IoT-enabled devices have generated huge volume of data [36]. The handling of massive amounts of structured, unstructured, and semi-structured data results requires special processing for such big data [37]. The current intensive data processing needs from IoT devices indicate a promising path of combining big data technologies specially aiming at IoT applications [38]. The massive amount of data needs high storage and high computing power and strong data analytics, as the traditional database systems are not able to store, process, and analyze such a huge amount of data [39]. The data generated from these devices are analyzed and used for the current and future decision-making processes.

The three primary determinants of big data, known as the “3Vs of big data," which are volume, velocity, and variety. The term "volume" refers to the massive amount of data collected, which causes datasets to be too vast for traditional database technology to handle. Larger data units, such as terabytes, petabytes, and exabytes, are used to describe this type of data. Velocity is the speed with which the data is generated, processed, and moved around in real time. whereas, The nature of data, whether it is structured or unstructured, is a source of variety. Big data analytics can handle massive amounts of structured, unstructured, and semi-structured data created by IoT devices in industrial equipment etc. Big data analytics can help companies produce and store and generate information from insight the data.

There are a few requirements that the data network related to the IoT needs to fulfill in order to support big data analytics. In the following, we present such necessary features.

**Connectivity**: Machine-to-machine (M2M) communication protocols that manage a large number of streams are the most common foundation for IoT services, and they benefit directly from cloud distributed storage and computation infrastructure. [40]. Secure connectivity for enormous data and analytics is the first and most basic requirement of the Internet of Things (IoT). Because of reliable connectivity, big data and analytics will be able to efficiently aggregate and integrate massive amounts of machine-generated sensor data [41].

**Storage**: In IoT, big data storage must be able to manage large amounts of unstructured data while also providing low latency for analytics. One problem is that there are numerous sources of IoT data, such as sensor data and social media, and they are all modeled differently utilizing different communication protocols and interfaces. Big data technology can help with data storage for IoT devices, but more solid solutions are needed.

**Quality of services**: Quality of service (QoS) refers to the capacity to ensure a specified level of performance for a data flow. The IoT network is responsible for providing the assurance of an effective transmission of data from the sources that generate large data. The quality of service (QoS) in an IoT network is critical for big data analytics [42].

**Real time analytics**: Information regarding IoT-connected objects is exchanged in real time, and it must be processed in real time. For most streaming data from web-enabled devices, big data analytics uses real-time queries to extract information fast, make choices, and interact with devices and people in real time [43].

**Benchmark**: Many corporations have begun to migrate their operations online utilizing IoT as a result of the rapid digitization of operations. Benchmarking is critical in this situation because it allows businesses to compare the quality of big data and analytics solutions [44].

### BENEFITS OF USING IOT AND BIG-DATA IN INDUSTRIES

For the industrial sector, there are numerous benefits of IoT and big data analytics. They provide information in a better way for the current and future decisions of the companies, using which the companies can build future business strategies and plans. Some of the benefits are listed below.

**Improve energy efficiency**: Energy is one of the biggest expenses in the industries, and industries are trying to reduce and save energy consumption. IoT with the help of big data is providing help to the industries in achieving that goal by providing the energy data on the devices level at real time. This can identify the under-performing devices in the network, and the energy management team can take the necessary actions to improve the energy wastage.

**Improved forecasting and predictive maintenance**: Using IoT and big data, automated alerts from the device provide useful information of the machine's maintenance, instead of waiting for the historic data. The staff concerned can know the machine's health in real time and plan the repair and replacement more efficiently, saving a lot of time.

**Improved product quality**: Product quality is the most important part of the industry. High-quality products improve customer satisfaction, sales, profit, and ultimately reduce waste. Using IoT, sensors can detect slight changes in the configuration, and big data analytics can make quick calculations and send alerts to the concerned staff. Thus, any defects can be fixed easily, thereby improving the product quality.

**Decrease downtime**: Using IoT, the performance of industrial machines is improved; this not only enhances the quality of the products but also improves the speed and performance of production, thus helping to complete the production on time without any issues.

**Quick accurate decisions**: The decision process in industries has improved using IoT. Using a machine's performance data, which, after collection by the sensor, is calculated by the big data, the managers take necessary steps to improve the organizational processes and overall productivity.

**Customer satisfaction**: The success of a company depends on the customer's satisfaction. If a company provides good products of high standards, the customers will recommend the same products to others as well, and the company business will be high. IoT and big data analytics together thus facilitate the company to develop high quality products.

## VI. SURVEY RESULTS AND DISCUSSION:

The survey conducted in this study contains reply from ten (10) big companies. The survey questionnaire contains twenty(20) questions. Table 1 and Figure3 show the survey questions and answers. The percentage of respondents with responses of "totally disagree" and "disagree" is almost negligible. A few of the company's experts were neutral to some of the questions, such as those on the lack of management awareness towards IEnMS and the creation of an energy management team by the energy manager. In general, most of the companies expressed the opinion of either "totally agreed" or "agreed" to most of the points, for example, the questions about whether the industries should use IEnMS using the latest technologies like IoT, big data and data analytics and whether this will facilitate the industries in terms of less energy consumption, efficient utilization of energy, reduced energy bills and costs and reductions in greenhouse gas emissions.

Figure 3. gives a pictorial representation of the results. Here, the purple color shows "totally agree"; light gray shows "agree"; light orange shows "neutral"; red line shows "disagree"; and blue shows "totally disagree" opinions of the industrial experts. Figure 3 shows company results of the survey questions. In Figure 3, In general, all the companies were more than 80% in agreement with the twenty questions that we have asked. The companies agreed about the importance of IEnMS in terms of energy efficiency, energy consumption, reduction in energy costs, and reductions in greenhouse emissions by utilizing renewable energy.

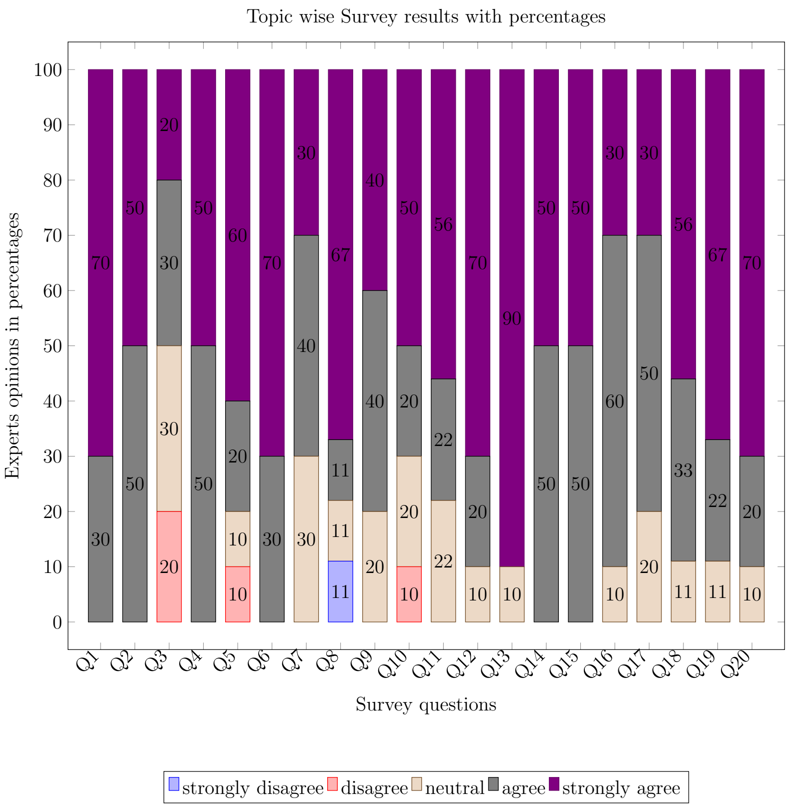



**FIGURE 3.** Topic wise survey results with percentages.

## VII. PROPOSED HIGH LEVEL ARCHITECTURE FOR IENMS

The huge amount of data generated by the IoT services creates opportunities to improve industrial services and customer values. For example, the data generated from IoT sensors can be analyzed in real time using big data that can help to take better presentation of information from the data and helps better future decisions and will result in continuous improvements within operations.

The proposed architecture is explained in Fig. 4. The process is divided into different phases, which are presented next.

**Initial Phase**: Data is collected from a variety of sources in the early phase, including machines, HVAC, data creation from renewable sources (solar, wind), lighting, CCTV, and a variety of other energy-consuming devices, using sensors and actuators. The vast amount of data generated by these devices is stored in the cloud at a minimum cost.

**Second Phase**: The second phase is data acquisition, in which the generated big data is stored in a shared distributed fault tolerant database depending on volume, velocity, and variety. The acquired data is then sent to the Hadoop cluster's master node(s). Because the data is acquired from a variety of heterogeneous devices, it may contain a variety of data formats and information, necessitating data preparation. Accurate and incomplete data are handled in data preparation, and incomplete data is either repaired or removed. The data collecting method is carried out using Flume. Flume's main job is to gather, combine, and send enormous amounts of data to the Hadoop master node.Flume stores the data it receives in a single or several channels.

**Third Phase**: The data is subsequently transmitted to an external Hadoop Distributed File System (HDFS) repository, where it is serialized and written in the desired format. Note that HDFS works by dividing large files into multiple blocks and stores these individual blocks on multiple data-nodes that are linked to master node. HDFS is generic enough that allows basically all types of data (structured, unstructured or semi structured data) to store. The serializers rearrange and alter the Flume data to fit the intended format.The data is

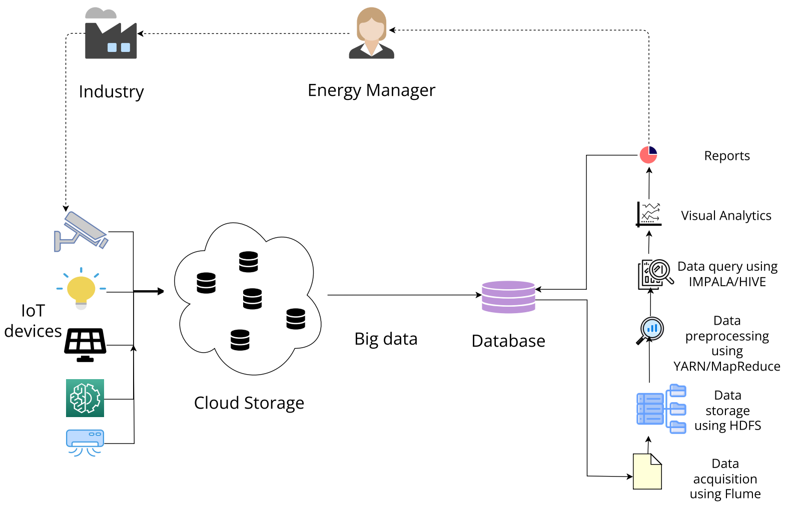



**FIGURE 4.** Data generation using IoT and processing with big data. The solid arrows mean data flow from one element to another; dashed line means possible information about the industrial operations based on the data reports from the IEnMS.

pre-processed, resulting in a unified picture of the data.For processing, the data is stored in several HDFS clusters. DataNodes make up the HDFS clusters.In the DataNodes, the actual data and file system meta data are kept. YARN analyzes data stored in HDFS; the two run on the same set of nodes, allowing jobs to be handled on nodes where EnM data is present.

**Fourth Phase**: This stage involves running SQL queries. Hive and Impala are two tools that may be used to run SQL queries on HDFS data. HIVE is used for data querying, to select, analyse, and to make calculations on the data of interest.

**Last Phase**: Data analytics is the final phase, in which the calculated data is shared with energy management, particularly the energy manager, to allow for better planning and decision-making in order to achieve efficient energy utilisation, reduced greenhouse gas emissions, increased machine efficiency, energy policy design, and improved energy planning, among other things. The tool used in Hadoop for data analytics is Scalable Advanced Massive Online Analysis (SAMOA), a distributed streaming machine learning framework that comprises of programming abstraction for distributed streaming algorithms for data mining and machine learning applications. Tableau is used for data visualisation (graphs, reports, and so on). Tableau is a popular tool for interactive data visualisation.

All in all, the expected outcome of the IEnMS process can be systematized into two different approaches, namely operational and strategic. At the operational level, the IEnMS serves as either an autonomous agent that is capable of intervening in the energy consumed throughout the targeted industrial process by, for example, an energy-centric schedule policy of operation or it can be used to monitor the energy demanded by different industrial tools in order to detect anomalies in their operation in terms of energy consumption. At strategic levels, the IEnMS may serve to support administrative policies targeting sustainability and energy efficiency by creating detailed data-driven reports that locate the most energy hungry or energy inefficient devices/processes that might be the target of future changes. It is nevertheless important to reinforce that our guidelines are general ones, which should be valid across different industries and cases; therefore, we avoid using particular cases and specific standards that would decrease the scope of our findings.

## VIII. CONCLUSIONS

This research highlights the importance of IEnMS and outlines its components in detail. The IEnMS describes the practices and processes of industrial energy management, which are typically thought of as supporting functions, as well as how they are carried out. We've addressed why IEnMS should use cutting-edge technologies like IoT, big data, and data analytics, as well as the benefits they provide to the industrial sector. To obtain the opinions of industrial experts about the importance of using the modern technologies in IEnMS, we conducted a detailed survey of large companies. The results show that most of the industrial experts are in favor of utilizing modern technologies in IEnMS. Based on the industrial experts' opinions, we have designed a theoretical framework for obtaining energy information using IEnMS and modern technologies. In this framework, the data from machines are collected using IoT devices and then transferred to a database from where the big data process and data analytics begin. The information generated from the data is finally send to the energy management expert (energy manager). Through this approach, industries are expected to improve their energy efficiency while reducing their energy consumption, costs, and greenhouse gas emissions. Moreover, the information retrieved from the data can be used for the current and future business decisions and for the maintenance of industrial machines. In particular, a survey carried out with energy experts from different companies indicate that gains in energy efficiency are expected to be achieve using IEnMS as a centralized entity that will support the company to take better strategic and operational decisions based on data, while enabling an energy-centric operation through specialized IoT devices, sensors and actuators. In future, the same study can be extended to include Power-to-Gas (PtG) technology that mainly convert the electric energy (in this case from renewable) into H2 using electrolysis process and later synthetic methane using the chemical reaction called methanation. These two (H2 and methane) can be store for longer period which can solve the long storage problem of renewable energy.

## REFERENCES

[1] A. Sa, P. Thollander, and E. Cagno, "Assessing the driving factors for energy management program adoption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 538–547, 2017.

[2] E. Pinero, "Iso 50001: setting the standard for industrial energy management," *Green Manufacturing News*, pp. 21–24, 2009.

[3] M. Wei, S. H. Hong, and M. Alam, "An iot-based energy-management platform for industrial facilities," *Applied energy*, vol. 164, pp. 607–619, 2016.

[4] H. M. Hussain, A. Narayanan, P. H. Nardelli, and Y. Yang, "What is energy internet? concepts, technologies, and future directions," *Ieee Access*, vol. 8, pp. 183 127–183 145, 2020.

[5] P. H. Nardelli, H. Alves, A. Pinomaa, S. Wahid, M. D. C. Tomé, A. Kosonen, F. Kühnlenz, A. Pouttu, and D. Carrillo, "Energy internet via packetized management: Enabling technologies and deployment challenges," *IEEE Access*, vol. 7, pp. 16 909–16 924, 2019.

[6] A. E. Nezhad, A. Rahimnejad, P. H. Nardelli, S. A. Gadsden, S. Sahoo, and F. Ghanavati, "A shrinking horizon model predictive controller for daily scheduling of home energy management systems," *IEEE Access*, 2022.

[7] D. P. Van Soest and E. H. Bulte, "Does the energy-efficiency paradox exist? technological progress and uncertainty," *Environmental and Resource Economics*, vol. 18, no. 1, pp. 101–112, 2001.

[8] S. T. Anderson and R. G. Newell, "Information programs for technology adoption: the case of energy-efficiency audits," *Resource and Energy Economics*, vol. 26, no. 1, pp. 27–50, 2004.

[9] J. Catarino, J. Henriques, and F. Egreja, "Portuguese sme toward energy efficiency improvement," *Energy Efficiency*, vol. 8, no. 5, pp. 995–1013, 2015.

[10] S. Sorrell, J. Schleich, S. Scott, E. O'Malley, F. Trace, U. Boede, K. Ostertag, and P. Radgen, "Reducing barriers to energy efficiency in public and private organizations," *Science and Policy Technology Research (SPRU), University of Sussex, Sussex, UK*, 2000.

[11] H. L. De Groot, E. T. Verhoef, and P. Nijkamp, "Energy saving by firms: decision-making, barriers and policies," *Energy Economics*, vol. 23, no. 6, pp. 717–740, 2001.

[12] J. Harris, J. Anderson, and W. Shafron, "Investment in energy efficiency: a survey of australian firms," *Energy Policy*, vol. 28, no. 12, pp. 867–876, 2000.

[13] P. Sandberg and M. Söderström, "Industrial energy efficiency: the need for investment decision support from a manager perspective," *Energy policy*, vol. 31, no. 15, pp. 1623–1634, 2003.

[14] M. Rudberg, M. Waldemarsson, and H. Lidestam, "Strategic perspectives on energy management: A case study in the process industry," *Applied Energy*, vol. 104, pp. 487–496, 2013.

[15] P. Thollander and J. Palm, "Industrial energy management decision making for improved energy efficiency—strategic system perspectives and situated action in combination," *Energies*, vol. 8, no. 6, pp. 5694–5703, 2015.

[16] ——, *Improving energy efficiency in industrial energy systems: An interdisciplinary perspective on barriers, energy audits, energy management, policies, and programs*. Springer Science & Business Media, 2012.

[17] P. Peura, "From malthus to sustainable energy—theoretical orientations to reforming the energy sector," *Renewable and Sustainable Energy Reviews*, vol. 19, pp. 309–327, 2013.

[18] B. Purwanggono, K. Ferastra, and A. Bachtiar, "Critical success factors evaluation of the iso 50001 energy management system implementation (case study: Pt. apac inti corpora, bawen, semarang indonesia)," in *IOP Conference Series: Materials Science and Engineering*, vol. 598, no. 1. IOP Publishing, 2019, p. 012115.

[19] M. Schulze, H. Nehler, M. Ottosson, and P. Thollander, "Energy management in industry–a systematic review of previous findings and an integrative conceptual framework," *Journal of cleaner production*, vol. 112, pp. 3692–3708, 2016.

[20] L. Sivill, J. Manninen, I. Hippinen, and P. Ahtila, "Success factors of energy management in energy-intensive industries: Development priority of energy performance measurement," *International journal of energy research*, vol. 37, no. 8, pp. 936–951, 2013.

[21] E. V. Tachmitzaki, E. A. Didaskalou, and D. A. Georgakellos, "Energy management practices' determinants in greek enterprises," *Sustainability*, vol. 12, no. 1, p. 133, 2020.

[22] A. Sa, P. Thollander, and M. Rafiee, "Industrial energy management systems and energy-related decision-making," *Energies*, vol. 11, no. 10, p. 2784, 2018.

[23] P. Thollander and M. Ottosson, "Energy management practices in swedish energy-intensive industries," *Journal of Cleaner Production*, vol. 18, no. 12, pp. 1125–1133, 2010.

[24] M. G. Rietbergen and K. Blok, "Setting smart targets for industrial energy use and industrial energy efficiency," *Energy Policy*, vol. 38, no. 8, pp. 4339–4354, 2010.

[25] R. Martin, M. Muûls, L. B. De Preux, and U. J. Wagner, "Anatomy of a paradox: Management practices, organizational structure and energy efficiency," *Journal of Environmental Economics and Management*, vol. 63, no. 2, pp. 208–223, 2012.

[26] S. Vasudevan and B. Higgins, "Strategic energy risk management for end users," *The Journal of Structured Finance*, vol. 10, no. 1, pp. 74–78, 2004.

[27] E. Abdelaziz, R. Saidur, and S. Mekhilef, "A review on energy saving strategies in industrial sector," *Renewable and sustainable energy reviews*, vol. 15, no. 1, pp. 150–168, 2011.

[28] R. W. Apeaning and P. Thollander, "Barriers to and driving forces for industrial energy efficiency improvements in african industries–a case study of ghana's largest industrial area," *Journal of Cleaner Production*, vol. 53, pp. 204–213, 2013.

[29] K. Bunse, M. Vodicka, P. Schönsleben, M. Brülhart, and F. O. Ernst, "Integrating energy efficiency performance in production management–gap analysis between industrial needs and scientific literature," *Journal of Cleaner Production*, vol. 19, no. 6-7, pp. 667–679, 2011.

[30] R. D. Peterson and C. K. Belt, "Elements of an energy management program," *Jom*, vol. 61, no. 4, pp. 19–24, 2009.

[31] D. Gordić, M. Babić, N. Jovičić, V. Šušteršič, D. Končalović, and D. Jelić, "Development of energy management system–case study of serbian car manufacturer," *Energy Conversion and Management*, vol. 51, no. 12, pp. 2783–2790, 2010.

[32] S. A. Ates and N. M. Durakbasa, "Evaluation of corporate energy management practices of energy intensive industries in turkey," *Energy*, vol. 45, no. 1, pp. 81–91, 2012.

[33] S. Suk, X. Liu, and K. Sudo, "A survey study of energy saving activities of industrial companies in the republic of korea," *Journal of Cleaner Production*, vol. 41, pp. 301–311, 2013.

[34] V. Blass, C. J. Corbett, M. A. Delmas, and S. Muthulingam, "Top management and the adoption of energy efficiency practices: Evidence from small and medium-sized manufacturing firms in the us," *Energy*, vol. 65, pp. 560–571, 2014.

[35] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I. A. T. Hashem, A. Siddiqa, and I. Yaqoob, "Big iot data analytics: architecture, opportunities, and open research challenges," *ieee access*, vol. 5, pp. 5247–5261, 2017.

[36] B. J. Gantz, D. Reinsel, and B. D. Shadows, "Big Data , Bigger Digital Shadow s , and Biggest Grow th in the Far East Executive Summary: A Universe of Opportunities and Challenges," *Idc*, vol. 2007, no. December 2012, pp. 1–16, 2012.

[37] K. Kambatla, G. Kollias, V. Kumar, and A. Grama, "Trends in big data analytics," *Journal of parallel and distributed computing*, vol. 74, no. 7, pp. 2561–2573, 2014.

[38] S. Choudhury, Q. Ye, M. Dong, and Q. Zhang, "IoT Big Data Analytics," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[39] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A. Hung Byers *et al.*, *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, 2011.

[40] G. Suciu, V. Suciu, A. Martian, R. Craciunescu, A. Vulpe, I. Marcu, S. Halunga, and O. Fratu, "Big data, internet of things and cloud convergence–an architecture for secure e-health applications," *Journal of medical systems*, vol. 39, no. 11, pp. 1–8, 2015.

[41] E. Al Nuaimi, H. Al Neyadi, N. Mohamed, and J. Al-Jaroodi, "Applications of big data to smart cities," *Journal of Internet Services and Applications*, vol. 6, no. 1, pp. 1–15, 2015.

[42] J. Jin, J. Gubbi, T. Luo, and M. Palaniswami, "Network architecture and qos issues in the internet of things for a smart city," in *2012 International Symposium on Communications and Information Technologies (ISCIT)*. IEEE, 2012, pp. 956–961.

[43] Z. Ding *et al.*, "Iot-statisticdb: a general statistical database cluster mechanism for big data analysis in the internet of things," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 535–543.

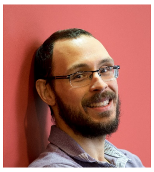
[44] A. Shukla, S. Chaturvedi, and Y. Simmhan, "Riotbench: An iot benchmark for distributed stream processing systems," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 21, p. e4257, 2017.


MEHAR ULLAH has B.S. in Information Technology from Iqra National University, Pakistan and Masters in Software Engineering from Lappeenranta-Lahti University of Technology (LUT), Finland, where He is currently a doctoral student, researcher. He is a member of the research group "Cyber-Physical Systems Group". His main research field is the role of IoT in cyber-physical systems and especially for industrial applications.

ARUN NARAYANAN (M'14) received his B.E. degree in Electrical Engineering from Visvesvaraya National Institute of Technology, Nagpur, India and M.Sc. in Energy Technology from Lappeenranta University of Technology (LUT), Finland, in 2002 and 2013, respectively. He subsequently completed his Ph.D. from the School of Energy Systems, LUT University, in 2019.

He is currently a Postdoctoral researcher with LUT University, Lappeenranta, Finland, in the research group "Cyber-Physical Systems Group." His research interests include renewable energy-based smart microgrids, electricity distribution and markets, demand-side management, energy management systems, and information and communications technology. He focuses on applying optimization, computational concepts, and artificial intelligence techniques to renewable electrical energy problems.

ANNIKA WOLFF has a BSc in Cognitive Science and a doctoral degree in knowledge modelling and HCI. She is a postdoctoral researcher at LUT University, Finland and a visiting Research Fellow at the Open University, UK. Her research is in the newly emerging field of human-data interaction, at the intersection between complex data, machine and human learning. Her research focuses on how people make sense of, interact with and design from complex data, such as derived from IoT and within smart cities. She is currently Co-I of H2020 (SWAFS) project ParCos on the Participatory Communication of Science, and conference planning chair for the European Society for Socially Embedded Technologies (EUSSET) as well as an associate editor for the International Journal of Human-Computer studies (IJHCS



PEDRO H. J. NARDELLI received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 2006 and 2008, respectively. In 2013, he received his doctoral degree from University of Oulu, Finland, and State University of Campinas following a dual degree agreement. He is currently Associate Professor (tenure track) in IoT in Energy Systems at LUT University, Finland, and holds a position of Academy of Finland Research Fellow with a project called Building the Energy Internet as a large-scale IoT-based cyber-physical system that manages the energy inventory of distribution grids as discretized packets via machine-type communications (EnergyNet). He leads the Cyber-Physical Systems Group at LUT and is Project Coordinator of the CHIST-ERA European consortium Framework for the Identification of Rare Events via Machine Learning and IoT Networks (FIREMAN). He is also Adjunct Professor at University of Oulu in the topic of "communications strategies and information processing in energy systems". His research focuses on wireless communications particularly applied in industrial automation and energy systems. He received a best paper award of IEEE PES Innovative Smart Grid Technologies Latin America 2019 in the track "Big Data and Internet of Things". He is also IEEE Senior Member. More information: https://sites.google.com/view/nardelli/

• • •

# Publication III

Ullah, M., Gutierrez-Rojas, D., Inkeri, E., Tynjälä, T., and Nardelli P. H. J.

**Operation of Power-to-X-Related Processes Based on Advanced Data-Driven Methods: A Comprehensive Review**

Reprinted with permission from
*MDPI energies*
Energies 2022, 15, 8118. https://doi.org/10.3390/en15218118

*Review*

# Operation of Power-to-X cogeneration plants based on advanced data-driven methods: A comprehensive review

**Mehar Ullah [1,‡] , Daniel Gutierrez-Rojas [1,‡] Eero Inkeri [1,‡] Tero Tynjälä [1] and Pedro H. J. Nardelli [1*]**

[1] School of Energy Systemes, Lappeenranta–Lahti University of Technology, Yliopistonkatu 34, 53850 Lappeenranta, Finland; [firstname.lastname]@lut.fi
* Correspondence: pedro.nardelli@lut.fi

**Abstract:** This study is a systematic analysis of selected research articles about the Power-to-x (P2X) sector. The relevance of this resides in the fact that most of the world energy is made from fossil fuels, which has been led to a huge amount of greenhouse gas emissions that are the source of global warming. One of the most supported actions against this issue is to produce and use the renewable energy resources, some of which are intermittent like solar and wind. This brings the need for large-scale, longer-period energy storage solutions. In this sense, P2X process chain could play this role: renewable energy can be converted into storable hydrogen, chemicals and fuels via electrolysis and subsequent synthesis with CO2. The focus of this study is to highlight the main contributions that are using the advanced data driven methods and latest technologies like Internet of Things (IoT), big data analytics and machine learning for the efficient operation of P2X cogeneration plants. We summarize our findings into different working architectures and illustrate it with a numerical result that employs machine learning model using the historic data and reduces the prediction error in a specific case of P2X.

**Keywords:** Power to X; IoT; big data; machine learning; electrolysis; methanation; synthetic gas

## 1. Introduction

This paper contains a comprehensive review of the operation of power-to-X (P2X) industrial plants following the conceptualization of cyber-physical systems introduced in [1]. It is critical to incorporate not only renewable energy sources, but also cutting-edge energy-efficiency technologies, in order to positively contribute to combating such an urgent issue and thus protecting current and future generations. The problem of climate change is inspiring a wave of helpful innovation. Exciting new concepts that could enhance our way of life are emerging. We cannot afford to waste energy if we are to successfully transition to a clean, dependable, and economical future energy system. Since the energy sector provides the energy needed by people and the economy. Europe is setting the pace in developing the upcoming energy system as it works to cut greenhouse gas emissions and transition to a sustainable society. The EU relies on energy that is safe, inexpensive, and environmentally sustainable, putting consumers first and fostering competition and growth. By doing so, it is estimated that, Cogeneration will serve as the backbone of a resilient, decentralized, and carbon-neutral European energy system by 2050, enabling industry and citizens in Europe to produce clean heat and power locally in a reliable, cost-effective, and efficient manner [2].

One of the primary strategies to quicken the energy transition will be P2X, especially when coupled with sector coupling. A few years ago, the idea of producing renewable, carbon-neutral fuels that absorb CO2 during production sounded like science fiction. This technology is now a fact. P2X collects CO2 from the atmosphere and combines it with green hydrogen to produce a variety of future fuels that are carbon-neutral. Sector coupling and P2X together pave the way for the decarbonization of numerous sectors. In a nutshell, it refers to tying together the energy-producing and consuming sectors, such as transportation

**Citation:** Ullah, M.; Gutierrez-Rojas, D.; Inkeri, E.; Tynjälä, T.; Nardelli, P. H. J. Operation of Power-to-X cogeneration plants based on advanced data-driven methods: A comprehensive review. *Energies* **2022**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

and industry, and converting them all to clean electricity from fossil fuels. The main forces behind sector coupling and P2X are accessible, renewable energy sources like wind and solar electricity. Since decarbonization and electrification go hand in hand, the development of new technologies that increase the viability of electrical power applications is naturally also essential to energy transition plans [3].

Belonging to P2X, also power-to-gas (P2G) systems could play a significant role in the future energy sector, which will rely more on intermittent renewable generation. One route for P2G is to produce methane (CH4), which can be used in place of natural gas, to provide long-term energy storage and variable load. The natural gas infrastructure already in place can be used to transport, store, and make use of produced CH4. P2G's primary processes are electrolysis to produce hydrogen (H2) from water and the conversion of hydrogen and carbon dioxide (CO2) to methane. A P2G system's crucial points also include the source, capture, and processing of the necessary CO2. When compared, for instance, to batteries, which are superior for short-term storage, P2G has the advantage of large capacity long-term storage. Utilizing extra low-emission energy is crucial to P2G since without it, the CO2 emissions of the generated CH4 tend to be excessively high in comparison to natural gas and bio gas [4].

The Internet of Things (IoT) is useful in this situation: IoT is able to collect and deliver this precise data from every section starting from power generation to a central place. Once there, it may be assessed before ideally permitting automatic control of the electrical grid. An IoT ecosystem is the end product. Due to the fast digitization of businesses, many organizations have started to shift their business online using IoT. Initially, these businesses were running very smoothly and the organizations where happy with the progress. However, with the growth in the businesses, they face challenges in storing and analyzing the huge amount of data connected through the IoT devices [5]. Finding solutions to those challenges requires some deep understanding of the problems. Big data and big-data analytics have received considerable attention from academia and many organizations. Big data and big-data analytics have found the solution of those challenges by using the big data and analytics platform [6].

With the use of machine learning (ML), which is a form of artificial intelligence (AI), software programs can predict outcomes more accurately, even without having to be explicitly instructed to do so. In order to forecast new output values, machine learning algorithms use historical data as input. Without being specifically programmed to do so, machine learning (ML), a subset of artificial intelligence (AI), enables software systems to improve their propensity to anticipate outcomes. Algorithms that use machine learning predict new values by using historical data as input. When predicting the likelihood of a specific outcome, machine learning refers to prediction as the result of an algorithm that has been trained on past data and applied to current data. The pace at which data is processed and evaluated is accelerated by machine learning. With very slight deployment adjustments, predictive analytics algorithms can now train on even larger data sets and do more in-depth research on a variety of aspects [7].

In this study we have focused on the data driven approaches for P2X technology. In the same study we have highlighted and explained the latest technologies like IoT , big data and analytics, machine learning , Artificial intelligence , P2G technology and its process. Later in this study we have presented a theoretical architecture that contains all these technologies together and shows their progress.

The rest of the paper is presented as section 2 contains the related study. Section 3 contains the basics and operations of the P2X cogenerations plant. In section 4 we have explained P2X as industrial cyber physical system and in the same section we have explained Internet of Things, artificial intelligence, meta-heuristics, machine learning. big data, and working of big data and at the end of this section we have highlighted requirements of big data analytics in IoT environment. Section 5 contains the theoretical architecture of the P2X cogeneration plant using the latest technologies discussed in 4 and in same section we have explained data processing for machine learning. Section 6 contains the discussion part of the paper.

## 2. Related study

The focus of this study is to highlight the importance of using the advanced data driven methods and latest technologies like IoT, Big data analytics and machine learning for the efficient utilizations and storage of the renewable energy used in the operation of power-to-X cogeneration plants like methanol synthesis using the advanced data driven methods. This study will contribute for the goals to reduce the greenhouse gas emissions. However, in this related study section.

A recent study [8] shows that by implementing a mixture of new technologies and apply put some investment can achieve the target of lowering the greenhouse gas emissions (CO2). This can be achieved by the combination of H2 (that can be generated by the electrolysis process by using the electricity generated by the renewable energy sources) and CO2 to generate synthetic natural gas (CH4). This process is called the methanation process and in this process H2 and CH4 is produced by using the electricity from the renewable sources [9]. The process is mainly called the P2X process and in this process, X here are the gases (H2 and CH4). CH4 produced in this process can be stored in the existing gas storage infrastructure and later can be used.

Hydrogen itself can be the final product of P2X. Certain amount, sometimes up to 20%, can be blended directly to the existing natural gas grid and utilized by the existing end use equipment. Larger share of H2 requires modifications or conversions to the equipment, due to different properties of H2 compared to natural gas. Additional challenge is the lower energy density of H2, which decreases the energy transport and buffer capacity (linepack swing) of the gas pipelines. [10,11]

During another study the author claims that P2X is a better option for the long-term storage of renewable energy sources as H2 and CH4 can be stored for long period of time as compared to battery storage and also reduce the amount of CO2 [12]. From the perspective of a Distribution System Operators DSO, P2G solves the problem of integrating RES both temporally and spatially due to insufficient distribution system capacities. Numerous research papers [12] have addressed the coordination of gas and electricity networks using P2G and other technologies. Several studies [13] evaluate the potentials of P2G on the transmission network level in countries like for example Germany. The authors claim a prediction of about 80 percentage of reduction of CO2 and about 110GW of energy by using the P2X technology by 2050 in the Northern side of the Germany due to high-speed wind and offshore capabilities.

There are some studies [14] that shows the requirements of predicting the future locations for the source of RES installations in different areas. They claim that by using the geographical information systems (GIS) services in those location can help in prediction of RES installation locations for generating the energy [15]. The above-mentioned studies are mainly focusing on the importance of RE and P2X technology for generating and storage of energy (H2, CH4) from the RES and also providing the importance of using the GIS services to identify the future RES locations for the installation. But, in our view point there is a research gap as none of the study has shown the data driven approaches for the collection of data from these renewable energy sources and also there is a gap for the future energy prediction using machine learning based of the future data driven from the methanation reactor.

In this study we have highlighted the importance of some data driven methods for collection of the data from the methanation reactor and later in the same study using the machine learning approach we have algorithm for the prediction of future energy cost for running the electrolysis process.

## 3. P2X Cogeneration Plants

### 3.1. Basics

The starting point of P2X is the production of hydrogen by water electrolysis, as hydrogen acts as the main energy carrier. There are three main technologies for the hydrogen production: alkaline electrolysis (AEL), polymer electrolyte membrane electrolysis (PEMEL), and solid oxide electrolysis (SOEL). AEL and PEM are already commercial, while SOEL is at pre-commercial phase. The main technical differences are related to pressure, temperature and dynamic operation. The operating temperature for AEL and PEMEL is about 50–80ºC, and 700–900ºC for SOEL. The maximum operating pressures for AEL, PEMEL and SOEL are 1–60, 4–76, and 10 bar, respectively. PEMEL is the most capable for low part load and fast transients, while AEL is also fast enough for grid frequency services, but has limited part load. SOEL requires more time for start-up and ramping, but is capable for wide load range. [16]

Hydrogen can be the final product of P2X, as it can be used directly as a fuel or raw material. Ammonia production and various chemical refining processes contribute to over 90% of the global hydrogen consumption of 73.9 Mt/a [17]. However, P2X can be extended by further processing bulk hydrogen to various products and materials.

While the most (69%) of the realized P2X projects in Europe do not process hydrogen further, methane is the second most used route (22%) [18]. The third one is methanol, which accounts for 6% of the projects.

Production of methane ($CH_4$) has been heavily studied for a long time, but the focus is changing form syngas to $CO_2$ methanation. In addition to hydrogen and $CO_2$, syngas contains also considerable amount of CO which changes the chemical process. Besides the different input composition, the general nature of process operation is changing from steady-state to transient operation of P2X. There are two main technological options for $CO_2$ methanation: biological and catalytic methanation. Methane has gained attention as it can be used to directly substitute natural gas with the existing grid and end use devices. Some commercial reactor concepts are already available. [19]

Other possible end products with high potential demand are methanol, dimethyl ether (DME), and Fischer-Tropcsh (FT) products. A significant benefit compared to methane is that the end products are in liquid form, thus increasing the energy density and usage potential. Similarly to the development of the methanation processes, the main challenge is the shift from well-known syngas process to $CO_2$-based process, and the transient operation of the plant. P2X processes for methanol and FT are more developed than for DME. [20]

The main structure for each of the P2X process are fairly similar at the upper level. As an example, a simplified process chain for methanol [21,22] is presented in figure 1. Renewable energy sources are the starting point, which provide the main energy input in form of electricity. The produced hydrogen and $CO_2$ are compressed to the operational pressure of the synthesis. Intermediate storage might be required for both hydrogen and $CO_2$, in which gases are stored above the synthesis pressure. After the synthesis, there are transportation and storage demand for the produced methanol.

In addition to the presented power and material flows, components consume and produce heat, thus benefit from heat integration [23]. Water and $CO_2$ sources are also needed for electrolysis and $CO_2$ capture [24].

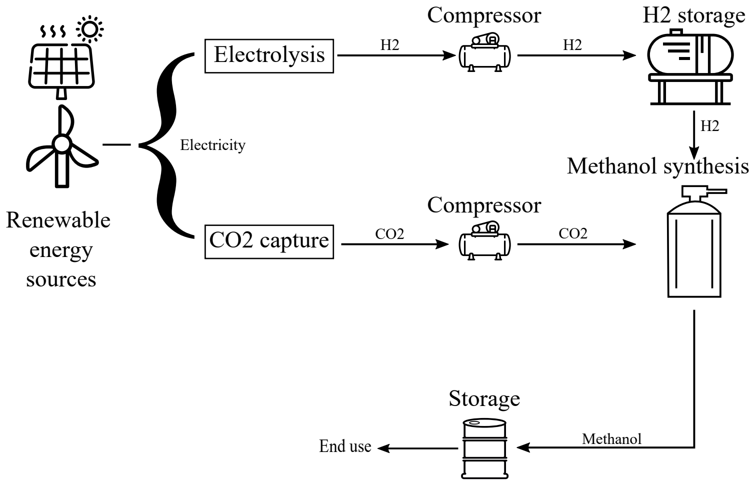



**Figure 1.** Working of the methanol synthesis process.

For more comprehensive information, reviews are available for electrolysis [16], methanation [19], and power-to-liquids (methanol, DME, FT) [20]. Examples of detailed process models can be found for alkaline water electrolysis [25], methanol synthesis [23], methanation [26].

### *3.2. Operation*

The operation of a P2X plant can be divided into two levels: scheduling and process control. Scheduling determines how the process interacts with the rest of the energy system, for example when power is used to generate hydrogen. The process control determines how the rest of the plant is operated with the given hydrogen feed. The target of scheduling is to obtain economic optimum, while the process control is used to keep to process within the technical limits, and optimize the technical performance. [27]

The scheduling of P2X systems is characterized by the variable energy input (wind and solar power) and price of the energy [21]. The material source of $CO_2$ is often considered rather constant [28]. Additional complexity is created due to different transient capabilities of system components. As an example, distillation is a crucial part of a methanol synthesis, and it is rather difficult to operate in a transient manner [28]. In contrast, electrolysers are able to operate in a very flexible manner [16].

In terms of the P2X process control, there are two extremes: (1) the load of the process follows strictly the inputs, or (2) the load is constant, and the variation of the inputs are leveled out by storage. Storage may be applied for one or several gases [29], or also for the electricity to produce hydrogen [30]. In any case, automation and control is required to start up, shut down, and maintain the process at certain set point.

For the scheduling, conventional optimization methods, such as linear programming (LP) or mixed-integer linear programming (MILP) have been used. These methods require usage of simplified physical models, that can be used as optimization constraints as defined by the optimization method. The process control require knowledge about the design and off-design performance, and the procedures of startup and load change. These two levels of operation, scheduling and process control, are interconnected.

Both scheduling and process control are effected by the environment in which the P2X plant is operated. The plant can be considered as standalone, so it does not affect to the rest of the energy system, and the only purpose of it is to make the end product with as low cost as possible. In this case, the plant just takes the resources and the prices as an input, which is a common assumption for techno-economic analyses [23]. Several end products may be considered, such as heat, oxygen, or grid services, but the revenue from them is considered only to decrease the production cost of the main end product [31].

Instead of minimizing the cost of the end product, the target could be maximizing of the revenues and the overall profitability of the P2X plant. This could include also operation as energy storage, by producing electricity back to grid when needed. The difference to the previous option is mainly in the way the costs and revenues are allocated.

Another option is that the plant acts as a part of the energy system, and the optimal operation is considered at the system level as in [32]. This way the total system cost can be minimized. However, the operation is not optimized from the point of view of a single P2X plant.

As the complete process models for the P2X plants can be very complicated, they can be computationally too heavy to be used for scheduling or real time operation control. Therefore, process models have been simplified to simple set of equations and constraints that can be implemented in the scheduling software, while losing some of the details of the physical model. Another option is to create a surrogate model [33] through machine learning, which is computationally lightweight but can still represent the physical behavior accurately. Cui et al. [34] used NARX model successfully for e-methanol plant. Tahkola [35] studied and compared four different machine learning methods for e-methane plant in Keras neural network: ARX, NARX, LSTM, and GRU. The resulting NRMSE was 1.94–3.60%. Shokry et al. [36] trained an AI model with two different knowledge of the chemical process: (1) only input and output signals were available as training data, and (2) mathematical model was available for creation of training data.

## 4. Data-driven operation of P2X Cogeneration Plants

### *4.1. P2X plant as an industrial cyber-physical system*

Processes that are constituted by logical decision making and physical relations are the definition of what is called cyber-physical systems (CPS) [1]. An important aspect under this concept is that every CPS is composed of 3 layers: physical layer, data layer and decision layer. In the case of P2X plants these layers are defined as following.

- **Physical layer:** This domain includes P2X plants that are used to physically perform the energy conversion. It also contains the measuring devices or sensors used to gather the information on analog variables. In the plant process, the variables from renewable energy sources, CO2 and H2 flow are examples.
- **Data layer:** This domain is where the analog of digital information is processed and converted into useful information about the plants variables. Contains relevant input information such as CO2 capture, spot price values of electricity, synthesis load, H2 storage, etc. that can later be injected into machine learning algorithms.
- **Decision layer:** In this domain is where the decision outcomes from the useful information from the data layer are involved. The decision can be either performed automatically by machines or by humans. In the process these decision making can be as the turn ON/OFF scheduling of the plant based on the predicting spot prices and the current state of the electrolyzer for energy storage.

The relationship between layers is close and data usually flows in loop from physical-data-decision-physical layer. The standardization of the plant and looking as a CPS serves the purposed to simplify the relation between the layers and find strong points within them to optimize the operation process. Usually communication between layers can also be a key factors as critical applications can be benefit from fast connections.

### *4.2. Optimization methods*

Optimization consist of the selection of the best element from a dominion space. It involves minimization or maximization of one or multiples variables of a function following a set of constraints. Optimization methods can be divided into two major categories: deterministic and stochastic methods (see Fig. 3). Deterministic methods can reach to a

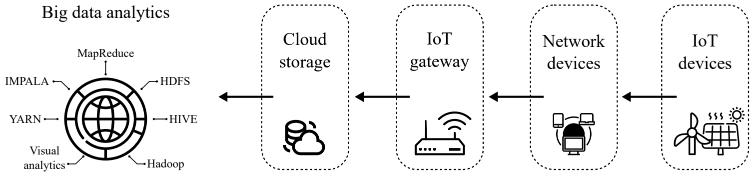



**Figure 2.** Working of IoT , Cloud and Big data

definite answer without uncertainty while stochastic methods reach to an approximate answer [37]. Moreover, another significant different between this major two classes, is that deterministic methods can take longer time to compute and for stochastic methods a wide range of different algorithms and programming tool-kits have been developed which makes it easier to adapt the objective function depending on the application. A big disadvantage of stochastic methods, is that due to its way to search the dominion space, it might get trapped in local minima (or maxima) when it comes to non-linear functions.

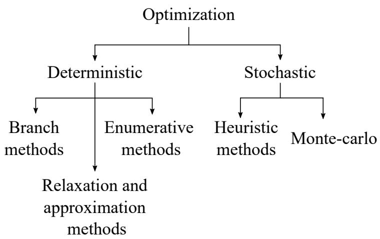



**Figure 3.** Overview of optimization methods (adapted from [38])

### *4.3. Internet of Things*

Internet of Things (IoT) one of the rapidly growing technology that is used to connect physical devices (objects) together using different communication techniques. IoT needs connectivity methods to ensure that the devices work, capture, analyze and manage the data accurately. For that purpose, connectivity should be flexible to meet the network performance needed for a wide range of IoT use cases, device types and applications. MQTT and HTTP are the important communication protocols that are used for the efficient connection and sharing of information [39] . The functionality and significance of IoT can be easily understand by studying the important building blocks of IoT; they are the components of IoT, which work together to deliver its functionality. There are six IoT building blocks that work together and provide functionality [40]. In the following, we will explain each of them in more details.

*Identification block:* This is the method that is used to identify devices in the network. There are two options for identifying the devices in the network object ID (name of device), and object address (the address of the device in the communication network) [41]. The main addressing methods of IoT objects are IPv6 and IPv4 [40].

*Sensing block:* Data is collected from the objects/devices using the sensor using the communication network and the collected data is then send to the cloud where it is analyzed. Actuators, i.e. hardware mechanical devices such as switches, are used in IoT platforms and operate in the opposite way to a sensor [40],[42],[43].

*Communication block:* The communication block contains heterogeneous objects that are used to exchange data and various services with each other and with the IoT platform. This block contains various IoT communication protocols like for example, MQTT and CoAP that are used to connect different objects to IoT and to send data from those connected objects to the management system. The devices like the sensors and other objects are connected to the Internet by communication technologies like for example, Zig-Bee, NFC, UWB, Wi-Fi, SigFox, and BLE [44], [40].

*Computation block:* This block contains two portions, software and hardware. The are plenty of hardware platforms that have been built to run IoT applications, for example, Raspberry PI, Intel Galileo, UDOO, Gadgeteer and Arduino. Similarly, there are many software platforms that are used to perform the functionalities of IoT. Operating system is the software the is running almost all the time during whole activation time of the device. The cloud platform is also a computational component of the IoT; it enables small objects to send data to the cloud, it facilitates big data processing in real time and helps the end user to obtain knowledge extracted from the big data [44],[40].

*Services block:* This is the block that provides the IoT application developers the starting point for IoT application. There are mainly four components of services block. The first one is *Identity related services* that can be divided into two parts, active and passive. Services that broadcast information and have a constant power or take power from the battery are active identity related services. Active identity related services can transmit or send information to another device. Passive identity related services have no power source and need some external device or mechanism to transmit its identity. Passive identity related services can only read information from devices. *Information aggregation services* refer to the actions of collecting data from sensors, processing that data, and transferring it to the IoT application for processing. *Collaborative aware services* use the data provided by the information aggregation services to make decisions and react accordingly. *Ubiquitous services* provide collaborative aware services anytime to anyone who needs it anywhere [42], [45],[46].

*Semantic Block:* IoT provides different services and for those services it needs knowledge, and for getting that knowledge in a better way, IoT uses different machines. Knowledge extraction can include finding and using resources, modeling information, and recognizing and analyzing data to reach some decision and provide the correct service. So, it can be claimed that the semantic block is the brain of the IoT [44],[40],[42].

### *4.4. Artificial intelligence*

Artificial intelligence (AI) has played a significant role in modern times, it has had a great impact in applications such as image recognition, medical applications, weather forecasting, etc that can be critical making positive changes for society [47]. Plenty of industrial and applications are designed everyday using AI methods, a important key factor for such acceptance is the simplicity and great accuracy to solve real-world problems. AI methods have 3 basic components, data, models and metrics. The amount of data and quality of the data are both important when it comes to AI, the methods uses historical information about the process to learn and model the behaviour between its elements and then be able to perform decision making in latter stages. AI models are based in connections and the more data there is, more re-validation of the model is performed. The models are the creation, training and testing of machine learning algorithms, they differ in each other by the different parametrization inside the layer that makes connections within the data. In order to evaluate the model and produce representative outcomes of the data based on logical decisions, metrics are use to estimate the differences between the model outcomes and data. Overall artificial intelligence has proved to be a powerful tool to deal with many of the problems that couldn't be solved or that were expensive to solved by providing approximate and enough accurate solutions. AI has taken part of those people who uses simple electronic equipment and its a great promise to potentiate industrial development. Integration of technologies in the frame of the Industry 4.0 would lead to increase efficiency in factories, increase techno-economical ratio and meet environmental urge of reducing carbon footprint [48].

### 4.5. Metaheuristics

Metaheuristics methods are search techniques that fall behind the heuristics definition without being problem dependant. Usually employed for optimization, metaheuristic techniques perform an intelligent search of the dominion space for a given function without the need of making a rigorous mathematical model [49]. Also metaheuristic algorithm have proven to be flexible and computationally cheap. In terms of the intelligent search provides by the metaheuristic, many principles or techniques have been proposed. Algorithms can be based on evolutionary programmed, trajectory, nature-inspired, ancient inspired, among others. For different application some might be more effective than others and the adaptation from one application to another is a simple task as long as the objective function, the boundaries and the restrictions are well defined. Moreover, similar to machine learning methods, the algorithms count with different parameters that can impact in how the space search is done. A good practice to approach is to tune the parameters by doing multiple rounds in order to find the optimal solution and check several times that the algorithms is not trapped in local minima. Metaheuristics are good candidate for industrial applications in P2X, they often require multi-objective optimization to solve energy scheduling problems.

### 4.6. Machine learning

Machine learning (ML) is an artificial intelligence discipline that enables machines to automatically learn from data and previous experiences while looking for patterns to make accurate predictions with minimal human involvement.

The working of machine learning can be explained in Figure 4. Initially the machine learning algorithm is trained with the training data-set (old data) in order to create a model. Here the machine learning algorithms are trained with the previous datasets. Once the machine learning model is trained, then the input data is provided to the trained machine learning algorithm for the calculations and future prediction. once the predictions are obtained from the input data, these predictions are evaluated and are matching with the actual results. If the predictions are close to the actual information then the decisions are made based on the predictions. In case the predictions are not close to the actual information(accuracy), then the machine learning algorithm is trained again with some more historic data (training data) and the same procedure is applied again to get a better predictions accuracy that is close to the actual information.

### 4.7. Working of IoT and big data

During the digitalization process, IoT devices are increasing and generating huge amount of data that needs to be processed in an efficient way to get information from it. Big data and IoT are the most needed technologies and the companies needs these two technologies to fulfil their business requirements and take effective decisions. IoT and big data are two different technologies, but with a passage of time and the needs for the efficient analysis of data, these two technologies becomes interrelated [50]. IoT is playing an important role in connecting and in communication between those devices and sharing of information between devices and with the IoT platform. The data generated from these devices is analyzed and used for the current and future decision-making processes. The number of IoT connecting devices are increasing, it is estimated that the number of IoT devices will reach to 20 billion in 2020 [51] and may be 50 billion in 2020 [52] and will generate huge volume of data of about 40 Zeta bytes in 2020 [53]. The data collected by these devices can be structured, unstructured and semi-structured data results in big-data [54], and handling, processing and storage of big data is not possible for the traditional data processing technologies and simple data bases. These problems data processing and storing can be handled effectively using the big data technologies to improve the development of IoT [55]. The massive amount of data needs high storage and high computing power and strong data analytics as the traditional databases systems are not able to store, process and analyze such a huge amount of data [56]. Cellular network like 5G is playing promising role for communication and data processing by providing good and fast communication between the connected objects. 5G is providing the ability to collect high amount of data, speedup the data analysis process, take out information from the data quickly which is used to make better predictions and speedup the decision process [57].

#### 4.7.1. Requirements of big data and analytics in IoT environment

IoT main goal is to connect various objects to the internet, collect data from those objects, analyze that data to get some meaningful information that helps the decision-makers to take good future business decisions. The need for big data and analytics to IoT has arisen as a result of the dramatic advancements in technology and business digitization over the past few years, as well as the growth in the number of devices connected to IoT. Big data and analytics have a high potential for extracting valuable information from the enormous amount of data and greatly improving the decision-making processes. Below is a description of the key requirements for big data and analytics in IoT, both functional and non-functional.

*Connectivity:* With the heterogeneous objects in the network and the numerous objects connected to the internet via sensors in a smart environment, connectivity in the IoT is largely ubiquitous. The provision of a dependable connectivity for big data and analytics is the first and most crucial requirement of the Internet of Things. The big data and analytics will have the chance to effectively combine and integrate the enormous amount of machine generated sensor data thanks to the reliable connectivity. Many of the objects in our environment can connect to the high-performance, high-computing infrastructure and support the IoT services using modern wireless networks like Wi-Fi and 4G/5G. [58].

*Storage:* The requirements of big data storage in IoT is to handle massive amount of unstructured data and provide low latency for analytics. Also, big data technology provides IoT an efficient data storage, processing and facilities to convert the massive amount of unstructured data into useful information which will provide a good foundation for better decision making. There are many sources of IoT data like sensors data, social media, smart phones etc. modeled in various ways and are using different communication protocols and interfaces. The IoT services are mostly based on machine-to-machine (M2M), communication protocols that are required to handle huge number of streams and is taking benefits directly from the cloud distributed storage and computing infrastructure [59].

*Quality of services:* Quality of service (QoS) refers to the capacity to guarantee a particular level of performance to the data flow. The IoT guarantees an efficient transfer of data from the sources that produce the big data, so the IoT network must be dependable and provide this guarantee. Big data and analytics rely heavily on the QoS in the IoT network. There are a variety of new networking technologies that can be used to build a dependable network, enable real-time event transfer for the Internet of Things, and enhance big data processing power[60].

*Real time analytics:* IoT is expanding quickly and making important improvements to streaming analytics and quick decision-making processes. Real-time information about the connected IoT objects is being communicated via IoT. Big data and analytics in IoT will accelerate the streaming process and extract the information as quickly as possible. Big data analytics performs real-time queries on the majority of the streaming data from web-enabled objects in order to obtain the information from it quickly, make decisions, and interact with the concerned devices and people in real-time. Big data uses an operational database for the streaming data. [61].

*Benchmark:* Many organizations have started moving their operations online using IoT as a result of the rapid digitization of businesses. Many of these businesses are having trouble storing and analyzing the enormous amount of data connected through IoT devices as a result of their expansion. It took a keen understanding of the issues to find the solutions to those problems. The academic community and many organizations are very interested in bid data and analytics. By utilizing the big data and analytics platform, big data and analytics have found a way to overcome those difficulties. In this situation, benchmarks

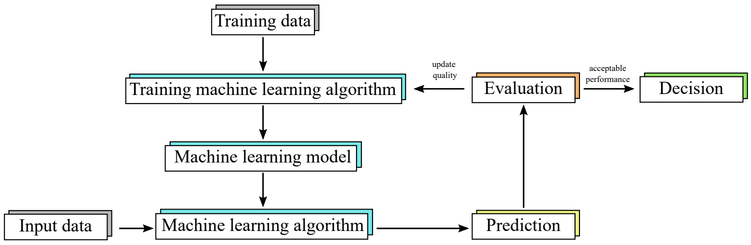



**Figure 4.** Working of Machine learning process

are crucial because they enable organizations to compare the effectiveness of big data and analytics solutions. [62].

## 5. Example: Theoretical Architecture for Deep learning based activation of electrolyzers

This section explains the advanced architecture of the P2G system by including the advanced data driven methods like, IoT, Big data and machine learning. The IoT enabled prediction system, is composed of four interlinked stages.

During the **first stage**, the IoT sensor devices are implemented in the whole infrastructure (Renewable energy sources, electrolysis process, hydrogen and methane storage etc.) of the P2G system for data collection. During this stage the data from various parts of the proposed architecture for example the data from the renewable source's solar panels, wind turbine, air pressure, atmospheric temperature, the amount of electricity generated by the energy sources, the amount of electricity provided to the electrolysis process and CO2 capture. How much H2 is generated and stored. How much H2 and CO2 is utilized to generate methanol and how much methanol is produced.

In the **second stage**, the generated data in stage one is stored in the cheap cloud storage for further processing.

In the **third stage** the huge amount of stored data is directed to the Big data analytics tools for further processing. The Big data consist of all the formats (structures, semi structures and unstructured) of data. The huge amount of data generated by the IoT services creates opportunities to improve industrial services and customer values. For example, the data generated from IoT sensors can be analyzed in real time using big data that can help to take better presentation of information from the data and helps better future decisions and will result in continuous improvements within operations. The Big data process is divided into different phases, which are presented next.

*Initial Phase*: The data acquisition stage is where big data generated during the previous phase is stored, depending on its volume, velocity, and variety, in a shared distributed fault-tolerant database. The master node of the Hadoop cluster is then sent the acquired data . Data preparation is necessary because the data may contain a variety of data formats and information because it is collected from a variety of heterogeneous devices. In data preparation, accurate and incomplete data are handled, and incomplete data is either fixed or removed. Using Flume, the data collection method is executed. The primary responsibility of Flume is to compile, combine, and send massive amounts of data to the Hadoop master node. Flume keeps track of the data it receives in one or more channels.

*Second Phase*: Following that, the data is sent to an outside Hadoop Distributed File System (HDFS) repository, where it is serialized and written in the required format. Be aware that HDFS stores individual blocks of large files on numerous data-nodes connected to the master node. It works by breaking up large files into multiple blocks. HDFS is sufficiently all-encompassing to enable the storage of essentially any type of data, whether it be structured, unstructured, or semi-structured.

To conform to the desired format, the serializers rearrange and modify the Flume data. Pre-processing the data yields a unified view of the data. The data is kept in various HDFS clusters for processing. The HDFS clusters are made up of DataNodes. The actual data and file system meta data are stored in the DataNodes. The two run on the same set of nodes, allowing jobs to be handled on nodes where the data is present. YARN analyzes data stored in HDFS.

*Third Phase*: SQL queries are executed during this phase. SQL queries can be executed on HDFS data using the tools Hive and Impala. HIVE is utilized for data querying, data selection, analysis, and computation on the pertinent data.

*Last Phase*:The last step, data analytics, involves sharing the calculated data to facilitate better planning and decision-making. Scalable Advanced Massive Online Analysis (SAMOA), a distributed streaming machine learning framework that includes programming abstraction for distributed streaming algorithms for data mining and machine learning applications, is the tool used in Hadoop for data analytics. To visualize data, use Tableau (graphs, reports, and so on). A well-liked tool for interactive data visualization is Tableau.

The **fourth stage** is the prediction stage. In this stage the calculated data from Big data storage is provided to the machine learning as training and input data-set. The rest of the machine learning process is explained in section 3.6 Figure 4.

*5.1. Data acquisition and communication network architecture*

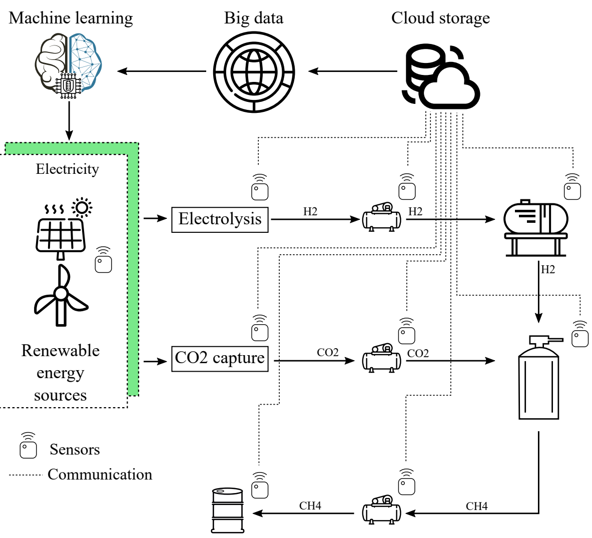



**Figure 5.** Working of the methanol synthesis.

*5.2. Data processing using machine learning*

Training data was produced with a Matlab model for 10.0 MW electrolysis, hydrogen storage, and methanol synthesis. Constant efficiency of 65% was assumed for the electrolysis, and stoichiometry is considered for the synthesis with a 100% conversion of input gases. A minimum part load of 80% was assumed for the synthesis, and it is not allowed to shut down the synthesis. The capacity of the hydrogen storage is 4.0 hours, determined based on the maximum hydrogen consumption of the synthesis. Feed rate of hydrogen to synthesis is a function of storage level and minimum part load of the synthesis. Capital costs of 750 €/kW, 500 €/kW and 600 €/kg were assumed for the electrolyser, synthesis, and hydrogen storage, respectively.

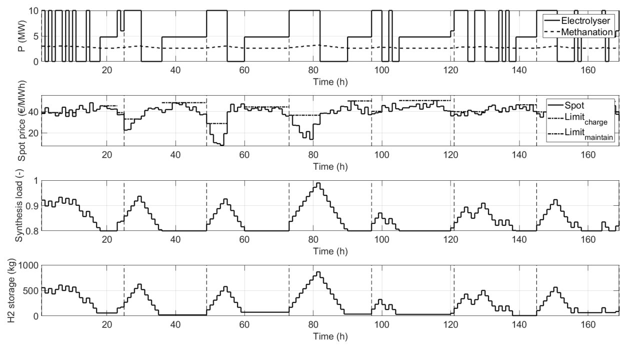



**Figure 6.** Optimization outcomes from methanol synthesis. a), Power b), Spot price, c) Synthesis load, d) H2 storage

An in-house, reduced brute force method was used to determine the operation of the electrolyser, in order to minimize the levelized cost of methane. An example result for a one week operation is presented in figure 6. As can be seen, the hydrogen storage is charged during the periods with cheap electricity, and discharged when electricity is expensive. For the rest of the periods, hydrogen is produced to only maintain the minimum part load of the synthesis.

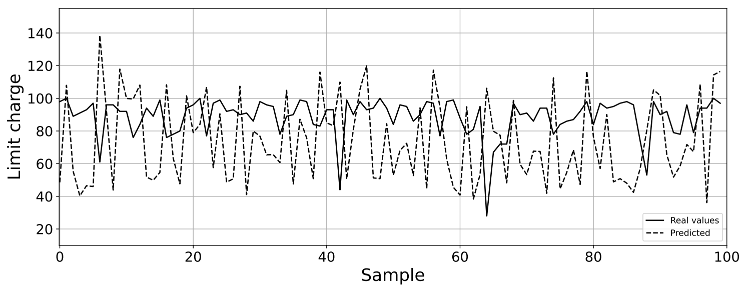



**Figure 7.** Regression model outcome compared with the limits obtained from optimization

The model is used to create the training data for the AI model. For this purpose, a set of 310690 days (24 h) were simulated. Each 24 h cycle was initiated with a random storage level, with random hourly electricity price of 0–100 €/MWh.

The training data was obtained from Matlab model, and it contains 310690 rows and 27 columns. Columns 1 to 24 are electricity prices, column 25 refers to the initial state of hydrogen buffer storage, column 26 is the price threshold to charge storage and column 27 is the price threshold to maintain minimum part load of the synthesis.

The outcome dataset was then used for training on a deep learning algorithm. The parameters of the algorithm were as follow:

- 5 hidden layer activated by ReLU function
- 1 output layer activated by linear function
- adam optimizer and mean squared error loss function
- 20 epoch with bach size of 1000

- 90% dataset for trainning, 10% for testing

In term of computational time, the model took about 185 *s* for training, about 9 *s* per epoch. The problem was treated as a regression and R2 score obtained was around -0.566. The comparison between curves obtained performing the regression testing and data can be seen in Fig. 7, they compare the upper limits of the price spot to produce hydrogen, based on the the 24 hours spot price and hydrogen buffer store. The variability of the data reflects on the model outcome as the trend is not easily followed.

## 6. Discussion

The current review study highlights the use of data driven methods and the use of latest technologies like IoT , big-data and machine learning in the operations of P2X plants by keeping the point that energy is important in all the fields of life.The electricity generated from the fossil fuels is expensive and can rise the greenhouse gas emissions that leads to the global warming. The study focus on the generation of the renewable electricity production by keeping in mind the long term storage of the renewable energy by using the latest technology P2X. The P2X technology can store the renewable energy for longer period by the generation of H2 from the electrolysis process and CH4 from the methanation process.

In this study we have discussed the importance of storing the renewable energy (green energy) for a longer period using the latest P2X technology. In the same study we have briefly explained and highlighted the importance of the latest technologies like IoT for data collection from the P2X plant , big-data for gathering the information from the huge amount of data generated by the IoT sensors and machine learning for the future prediction of information based of the data-set from the big-data sources.

Long term storage of the renewable energy from various sources like solar, wind etc. is a hard task as batteries are not able to store the energy for a longer period. To overcome the aforementioned problem, a lot of research has been done in the recent years about P2X technology and specially for P2G in the field of energy sector to store the energy in liquid form for longer period using the already build storage infrastructure of H2 and CH4.

The value of data is critical in the digital transformation era to support the monitoring, diagnosis, prediction, planning, and optimization of shop floor assets. The collection, transmission, storage, and analysis of the massive amount of available data are critical issues for this purpose.

This paper describes how data is gathered, distributed and automatically in P2X technology and introduces existing technologies for implementing distributed data collection systems. P2G technology has enormous potential for electric grid load balancing, renewable integration, and the hydrogen economy by connecting the power and gas vectors in a flexible utility network capable of storing energy from multiple sources and using it for a variety of applications. It must be made abundantly clear that RES-based energy systems, which gradually phase out fossil fuels, require dependable long-term storage (P2G fulfilling the required criteria only).By producing H2 as an intermediate product, converting energy through electrolysis, and storing the fuel until it has a high power consumption, when it will be converted back into electricity or used for other purposes, P2G differs from other storage technologies. A further step is considered to produce synthetic natural gas (SNG) to be injected into the gas grid, which has advantages over hydrogen or electricity in terms of storage, transport, and uses. It enables cross-sectoral integration of surplus, low-value renewable energy in energy-demanding sectors like transportation or industry, facilitating further decarbonization of these industries while simultaneously opening up new sources of system flexibility in the power sector.

According to the findings, P2G schemes are excellent candidates for changing the energy system to one that is more sustainable, at least while the transition is taking place. When non-adjustable power generation temporarily exceeds the loads or if demand cannot be met by the generation capacity, P2G is then continuously used. The potential for storage (residual loads) in both transitional and long-term scenarios can be seen in a yearly model on an hourly basis. It can be added without restrictions to the gas distribution network, which has significant technical potential for storage, transport, and uses, if it is synthesized properly as shown here (i.e., 99% yield, 95% CH4).

If conditions are changed so that the energy systems serve the technology, P2G appears to be a promising strategy in the long run. On the other hand, such strategies are likely to be the only ones capable of addressing transition process to truly sustainable technologies, where base load utilities can be operated more consistently and variable power capacities (RES) are utilized at their maximum possible energy contribution. Concerning the environmental component, there are essentially no alternatives to the integration of a very large fraction of renewable in the energy system, where P2G will likely be unavoidable even as technology must improve efficiencies and short-term costs; being technically feasible, it has the potential to manage systems formed by clean power sources that are converted into fuels (hydrogen and methane), whereas RES penetration requires balancing power and seasonal storage of electricity which can be expensive and time-consuming.

**Author Contributions:** Conceptualization, all; methodology, all; software, D.G:, E.I.; validation, D.G., E.I.; formal analysis, M.U., E.I.; investigation, M.U., D.G., E.I.; resources, X.X.; data curation, X.X.; writing—original draft preparation, all; writing—review and editing, T.T., P.N.; visualization, M.U., D.G.; supervision, T.T., P.N..; project administration, T.T.; funding acquisition,T.T., P.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is partly supported by (1) Academy of Finland via: (a) FIREMAN consortium n.326270 as part of CHIST-ERA grant CHIST-ERA-17-BDSI-003, (b) EnergyNet Fellowship n.321265/n.328869/n.352654, and (c) X-SDEN project n.349965; and Business Finland via HYGCEL project.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable

**Acknowledgments:** The authors gratefully acknowledge the public financing of Business Finland for the 'HYGCEL' project

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| P2X | Power-to-X |
| P2G | Power-to-gas |
| IoT | Internet of Things |
| ML | Machine learning |
| SNG | Syntatic natural gas |
| H2 | Hydrogen |
| CH4 | Methane |
| SAMOA | Scalable advanced massive online analysis |
| HDFS | Hadoop distributed file system |
| SQL | Sequential query language |
| QoS | Quality of service |
| AI | Artificial intelligence |
| CPS | Cyber physical system |
| CO2 | Carbon dioxide |
| LP | Linear programming |
| MILP | Mixed integer linear programming |
| GIS | Graphical information system |
| DSO | Distribution system operator |

## References

1. Nardelli, P.H. *Cyber-physical Systems: Theory, Methodology, and Applications*; John Wiley & Sons, 2022.

2. Child, M.; Kemfert, C.; Bogdanov, D.; Breyer, C. Flexible electricity generation, grid exchange and storage for the transition to a 100% renewable energy system in Europe. *Renewable energy* **2019**, *139*, 80–101.
3. Ozturk, M.; Dincer, I. A comprehensive review on power-to-gas with hydrogen options for cleaner applications. *International Journal of Hydrogen Energy* **2021**, *46*, 31511–31522.
4. Inkeri, E.; Tynjälä, T.; Karjunen, H. Significance of methanation reactor dynamics on the annual efficiency of power-to-gas-system. *Renewable Energy* **2021**, *163*, 1113–1126.
5. Ullah, M.; Nardelli, P.H.; Wolff, A.; Smolander, K. Twenty-one key factors to choose an iot platform: Theoretical framework and its applications. *IEEE Internet of Things Journal* **2020**, *7*, 10111–10119.
6. Ullah, M.; Narayanan, A.; Wolff, A.; Nardelli, P. Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing. In Proceedings of the 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2021, pp. 956–961.
7. Kanade Vijay. What is machine learning? Understanding types & applications, 2022.
8. Giehl, J.; Sudhaus, T.; Kurre, A.; Mikulicz-Radecki, F.v.; Hollnagel, J.; Wacker, M.; Himmel, J.; Müller-Kichenbauer, J. Modelling the impact of the energy transition on gas distribution networks in Germany. *Energy Strategy Reviews* **2021**, *38*, 100751.
9. Schiebahn, S.; Grube, T.; Robinius, M.; Tietze, V.; Kumar, B.; Stolten, D. Power to gas: Technological overview, systems analysis and economic assessment for a case study in Germany. *International journal of hydrogen energy* **2015**, *40*, 4285–4294.
10. Klatzer, T.; Bachhiesl, U.; Wogrin, S. State-of-the-art expansion planning of integrated power, natural gas, and hydrogen systems. *International Journal of Hydrogen Energy* **2022**, *47*, 20585–20603. https://doi.org/https://doi.org/10.1016/j.ijhydene.2022.04.293.
11. Quarton, C.J.; Samsatli, S. Should we inject hydrogen into gas grids? Practicalities and whole-system value chain optimisation. *Applied Energy* **2020**, *275*, 115172. https://doi.org/https://doi.org/10.1016/j.apenergy.2020.115172.
12. Farrokhifar, M.; Nie, Y.; Pozo, D. Energy systems planning: A survey on models for integrated power and natural gas networks coordination. *Applied Energy* **2020**, *262*, 114567.
13. Haumaier, J.; Hauser, P.; Hobbie, H.; Möst, D. Grünes Gas für die Gaswirtschaft–Regionale Power-to-Gas-Potentiale aus Onshore-Windenergie in Deutschland. *Zeitschrift für Energiewirtschaft* **2020**, *44*, 61–83.
14. Pick, J.B.; Perry, M.; Rosales, J. Application of GIS to Support Regional Policy for Development of Renewable Energy in Southern California: An Exploratory Case Study Analysis **2014**.
15. Henni, S.; Staudt, P.; Kandiah, B.; Weinhardt, C. Infrastructural coupling of the electricity and gas distribution grid to reduce renewable energy curtailment. *Applied Energy* **2021**, *288*, 116597.
16. Buttler, A.; Spliethoff, H. Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review. *Renewable and Sustainable Energy Reviews* **2018**, *82*, 2440–2454. https://doi.org/https://doi.org/10.1016/j.rser.2017.09.003.
17. IEA. The Future of Hydrogen, 2019.
18. Wulf, C.; Zapp, P.; Schreiber, A. Review of Power-to-X Demonstration Projects in Europe. *Frontiers in Energy Research* **2020**, *8*. https://doi.org/10.3389/fenrg.2020.00191.
19. Rönsch, S.; Schneider, J.; Matthischke, S.; Schlüter, M.; Götz, M.; Lefebvre, J.; Prabhakaran, P.; Bajohr, S. Review on methanation – From fundamentals to current projects. *Fuel* **2016**, *166*, 276–296. https://doi.org/https://doi.org/10.1016/j.fuel.2015.10.111.
20. Dieterich, V.; Buttler, A.; Hanel, A.; Spliethoff, H.; Fendt, S. Power-to-liquid via synthesis of methanol, DME or Fischer–Tropsch-fuels: a review. *Energy Environ. Sci.* **2020**, *13*, 3207–3252. https://doi.org/10.1039/D0EE01187H.
21. Chen, C.; Yang, A.; Bañares-Alcántara, R. Renewable methanol production: Understanding the interplay between storage sizing, renewable mix and dispatchable energy price. *Advances in Applied Energy* **2021**, *2*, 100021. https://doi.org/https://doi.org/10.1016/j.adapen.2021.100021.
22. Schorn, F.; Breuer, J.L.; Samsun, R.C.; Schnorbus, T.; Heuser, B.; Peters, R.; Stolten, D. Methanol as a renewable energy carrier: An assessment of production and transportation costs for selected global locations. *Advances in Applied Energy* **2021**, *3*, 100050. https://doi.org/https://doi.org/10.1016/j.adapen.2021.100050.
23. Battaglia, P.; Buffo, G.; Ferrero, D.; Santarelli, M.; Lanzini, A. Methanol synthesis through CO2 capture and hydrogenation: Thermal integration, energy performance and techno-economic assessment. *Journal of CO2 Utilization* **2021**, *44*, 101407. https://doi.org/https://doi.org/10.1016/j.jcou.2020.101407.
24. Chen, C.; Lu, Y.; Banares-Alcantara, R. Direct and indirect electrification of chemical industry using methanol production as a case study. *Applied Energy* **2019**, *243*, 71–90. https://doi.org/https://doi.org/10.1016/j.apenergy.2019.03.184.
25. Sakas, G.; Ibáñez-Rioja, A.; Ruuskanen, V.; Kosonen, A.; Ahola, J.; Bergmann, O. Dynamic energy and mass balance model for an industrial alkaline water electrolyzer plant process. *International Journal of Hydrogen Energy* **2022**, *47*, 4328–4345. https://doi.org/https://doi.org/10.1016/j.ijhydene.2021.11.126.
26. Fache, A.; Marias, F.; Guerré, V.; Palmade, S. Optimization of fixed-bed methanation reactors: Safe and efficient operation under transient and steady-state conditions. *Chemical Engineering Science* **2018**, *192*, 1124–1137. https://doi.org/https://doi.org/10.1016/j.ces.2018.08.044.
27. Zheng, Y.; You, S.; Li, X.; Bindner, H.W.; Münster, M. Data-driven robust optimization for optimal scheduling of power to methanol. *Energy Conversion and Management* **2022**, *256*, 115338. https://doi.org/https://doi.org/10.1016/j.enconman.2022.115338.

28. Chen, C.; Yang, A. Power-to-methanol: The role of process flexibility in the integration of variable renewable energy into chemical production. *Energy Conversion and Management* **2021**, *228*, 113673. https://doi.org/https://doi.org/10.1016/j.enconman.2020.113673.
29. Gorre, J.; Ruoss, F.; Karjunen, H.; Schaffert, J.; Tynjälä, T. Cost benefits of optimizing hydrogen storage and methanation capacities for Power-to-Gas plants in dynamic operation. *Applied Energy* **2020**, *257*, 113967. https://doi.org/https://doi.org/10.1016/j.apenergy.2019.113967.
30. Huesman, A. Integration of operation and design of solar fuel plants: A carbon dioxide to methanol case study. *Computers & Chemical Engineering* **2020**, *140*, 106836. https://doi.org/https://doi.org/10.1016/j.compchemeng.2020.106836.
31. Parra, D.; Patel, M.K. Techno-economic implications of the electrolyser technology and size for power-to-gas systems. *International Journal of Hydrogen Energy* **2016**, *41*, 3748–3761. https://doi.org/https://doi.org/10.1016/j.ijhydene.2015.12.160.
32. Bogdanov, D.; Gulagi, A.; Fasihi, M.; Breyer, C. Full energy sector transition towards 100Integrating power, heat, transport and industry sectors including desalination. *Applied Energy* **2021**, *283*, 116273. https://doi.org/https://doi.org/10.1016/j.apenergy.2020.116273.
33. McBride, K.; Sundmacher, K. Overview of Surrogate Modeling in Chemical Process Engineering. *Chemie Ingenieur Technik* **2018**, *91*, 228–239, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/cite.201800091]. https://doi.org/https://doi.org/10.1002/cite.201800091.
34. Cui, X.; Kær, S.K.; Nielsen, M.P. Energy analysis and surrogate modeling for the green methanol production under dynamic operating conditions. *Fuel* **2022**, *307*, 121924. https://doi.org/https://doi.org/10.1016/j.fuel.2021.121924.
35. Tahkola, M. Developing dynamic machine learning surrogate models of physics-based industrial process simulation models. PhD thesis, University of Oulu, Finland, 2019.
36. Shokry, A.; Baraldi, P.; Zio, E.; Espuña, A. Dynamic Surrogate Modeling for Multistep-ahead Prediction of Multivariate Nonlinear Chemical Processes. *Industrial & Engineering Chemistry Research* **2020**, *59*, 15634–15655, [https://doi.org/10.1021/acs.iecr.0c00729]. https://doi.org/10.1021/acs.iecr.0c00729.
37. Bangert, P. *Optimization for industrial problems*; Springer: Heidelberg ; New York, 2012. OCLC: ocn766342904.
38. Koopialipoor, M.; Noorbakhsh, A. Applications of Artificial Intelligence Techniques in Optimizing Drilling. In *Emerging Trends in Mechatronics*; Azizi, A., Ed.; IntechOpen: Rijeka, 2020; chapter 6. https://doi.org/10.5772/intechopen.85398.
39. Ullah, M.; Smolander, K. Highlighting the Key Factors of an IoT Platform. *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* **2019**, pp. 901–906. https://doi.org/10.23919/mipro.2019.8756748.
40. Al-Fuqaha, A.; Guizani, M.; Mohammadi, M.; Aledhari, M.; Ayyash, M. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE communications surveys & tutorials* **2015**, *17*, 2347–2376.
41. Koshizuka, N.; Sakamura, K. Ubiquitous ID: standards for ubiquitous computing and the internet of things. *IEEE Pervasive Computing* **2010**, *9*, 98–101.
42. Song, Y. Security in Internet of Things, 2013.
43. Sukanya Mandal. Internet of Things (IoT) - Part 2 (Building Blocks & Architecture), 2016.
44. Hejazi, H.; Rajab, H.; Cinkler, T.; Lengyel, L. Survey of platforms for massive IoT. *2018 IEEE International Conference on Future IoT Technologies, Future IoT 2018* **2018**, *2018-Janua*, 1–8. https://doi.org/10.1109/FIOT.2018.8325598.
45. Gigli, M.; Koo, S. Internet of Things: Services and Applications Categorization. *Advances in Internet of Things* **2011**, *01*, 27–31. https://doi.org/10.4236/ait.2011.12004.
46. Mohammed, F.H.; Esmail, R. Survey on IoT Services: Classifications and Applications. *International Journal of Science and Research* **2015**, *4*, 2124–2127.
47. Palhares, R.M.; Yuan, Y.; Wang, Q. Artificial Intelligence in Industrial Systems. *IEEE Transactions on Industrial Electronics* **2019**, *66*, 9636–9640. https://doi.org/10.1109/TIE.2019.2916709.
48. Vaddadi, S.; Srinivas, V.; Reddy, N.A.; H, G.; D, R.; Devipriya, A. Factory Inventory Automation using Industry 4.0 Technologies. In Proceedings of the 2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET), 2022, pp. 734–738. https://doi.org/10.1109/GlobConET53749.2022.9872416.
49. Gutierrez Rojas, D.; Lopez Lezama, J.; Villa, W. Metaheuristic Techniques Applied to the Optimal Reactive Power Dispatch: a Review. *IEEE Latin America Transactions* **2016**, *14*, 2253–2263. https://doi.org/10.1109/TLA.2016.7530421.
50. Marjani, M.; Nasaruddin, F.; Gani, A.; Karim, A.; Hashem, I.A.T.; Siddiqa, A.; Yaqoob, I. Big IoT data analytics: architecture, opportunities, and open research challenges. *ieee access* **2017**, *5*, 5247–5261.
51. Qi, Q.; Tao, F. Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *Ieee Access* **2018**, *6*, 3585–3593.
52. Ahmed, E.; Yaqoob, I.; Hashem, I.A.T.; Khan, I.; Ahmed, A.I.A.; Imran, M.; Vasilakos, A.V. The role of big data analytics in Internet of Things. *Computer Networks* **2017**, *129*, 459–471.
53. Gantz, B.J.; Reinsel, D.; Shadows, B.D. Big Data , Bigger Digital Shadow s , and Biggest Grow th in the Far East Executive Summary: A Universe of Opportunities and Challenges. *Idc* **2012**, *2007*, 1–16.
54. Kambatla, K.; Kollias, G.; Kumar, V.; Grama, A. Trends in big data analytics. *Journal of parallel and distributed computing* **2014**, *74*, 2561–2573.

55. Choudhury, S.; Ye, Q.; Dong, M.; Zhang, Q. IoT Big Data Analytics. *Wireless Communications and Mobile Computing* **2019**, *2019*. https://doi.org/10.1155/2019/9245392.
56. James Manyika.; Michael Chui. Big data: The next frontier for innovation, competition, and productivity | McKinsey, 2011.
57. Hui, H.; Ding, Y.; Shi, Q.; Li, F.; Song, Y.; Yan, J. 5G network-based Internet of Things for demand response in smart grid: A survey on application potential. *Applied Energy* **2020**, *257*, 113972.
58. Al Nuaimi, E.; Al Neyadi, H.; Mohamed, N.; Al-Jaroodi, J. Applications of big data to smart cities. *Journal of Internet Services and Applications* **2015**, *6*, 1–15.
59. Suciu, G.; Suciu, V.; Martian, A.; Craciunescu, R.; Vulpe, A.; Marcu, I.; Halunga, S.; Fratu, O. Big data, internet of things and cloud convergence–an architecture for secure e-health applications. *Journal of medical systems* **2015**, *39*, 1–8.
60. Jin, J.; Gubbi, J.; Luo, T.; Palaniswami, M. Network architecture and QoS issues in the internet of things for a smart city. In Proceedings of the 2012 International Symposium on Communications and Information Technologies (ISCIT). IEEE, 2012, pp. 956–961.
61. Ding, Z.; Gao, X.; Xu, J.; Wu, H. IOT-StatisticDB: a general statistical database cluster mechanism for big data analysis in the internet of things. In Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing. IEEE, 2013, pp. 535–543.
62. Shukla, A.; Chaturvedi, S.; Simmhan, Y. Riotbench: An iot benchmark for distributed stream processing systems. *Concurrency and Computation: Practice and Experience* **2017**, *29*, e4257.

# Publication IV

Ullah, M., Gopalraj, S. K., Gutierrez-Rojas, D., Nardelli, P., and Kärki, T.

**IoT framework and requirement for intelligent industrial pyrolysis process to recycle CFRP composite wastes: application study**

Reprinted with permission from
*Springer*
pp. 275–282. https://doi.org/10.1007/978-3-031-18641-7 26

# IoT framework and requirement for intelligent industrial pyrolysis process to recycle CFRP composite wastes: application study

Mehar Ullah[1[0000-0003-2405-8998]], Sankar Karuppannan Gopalraj[2], Daniel Gutierrez-Rojas[1], Pedro Nardelli[1] and Timo Kärki[2]

[1] School of Energy Systems, LUT University, Yliopistonkatu 34, 53850 Lappeenranta, Finland
[2] Fiber Composite Laboratory, Department of Mechanical Engineering, LUT University, P.O. Box 20, 53850 Lappeenranta, Finland.
Mehar.Ullah@lut.fi

**Abstract.** The cumulating carbon fiber-reinforced polymer (CFRP) composite wastes need to be disposed efficiently. So far, the most effective thermal-based recycling technique, namely pyrolysis, has grown exponentially towards industrial scaling in developed countries such as the UK and Germany. Typically, even the slightest mistakes can cause unfavorable results and delays in workflow within such a massive operating environment (e.g., > 1 tonnes/day operating capacity). The existing semi-automated and, in some cases, fully automated plants should be continuously updated to resemble the varying classes and volume of the CFRP composite wastes. To overcome such research gaps and imprecise manual errors, Internet-of-Things (IoT) based framework is proposed. This paper studies the theoretical implementation of an IoT-based framework into the pyrolysis process to recycle CFRP composite waste to manage the process based on the principles of cyber-physical systems. The proposed framework consists of sensors and actuators that will be used to collect the data and communicate with a central management constructed as a platform that will articulate and manipulate data to satisfy the requirements of the recycling process, computationally modeled through logical relations between physical entities. In this case, the management unit can be either controllable or monitored remotely to increase the operation time of the plant. Our objective is to propose a scalable method to improve the recycling process, which will also help future decision-making in handling recycled carbon fiber. Specifically, this study will go beyond the state-of-the-art in the field by (i) automatically calculate the mass of the waste and adjust the operating time, temperature, atmospheric pressure, and inert gas flow (if needed), (ii) regenerating heat so that after the first batch is recycled, the resin high in calorific value will be burned and will be releasing energy, whose generated heat needs to be trapped inside the furnace and then regenerated into the system, and (iii) decrease energy consumption and fasten the process flow time. In summary, the proposed framework aims to provide a user-friendly control and temperature monitoring that can increase the overall efficiency of the process and avoid possible process shut down or even char formation by a controlled atmosphere in the pyrolytic reactor.

**Keywords:** IoT Framework, Industrial-scale Pyrolysis, Recycling Carbon Fiber.

## 1 Introduction

Carbon fiber-reinforced polymer (CFRP) composites have been exponentially used in high-performance applications for decades. The composites have high mechanical properties for a lower weight ratio making them capable of replacing traditional metals in lightweight applications. However, CFRP composites employed 20 years ago have now reached their end-of-life (EoL) and raised a significant question about their disposal routes. So far, 62000 tonnes of CFRP composite wastes have been cumulating each year, and the forecast predicts that the amount could increase up to 90,000+ tonnes/year if not disposed properly. At the same time, the annual demand for virgin CFRP composites also expected to be increased from 72,000 to 140,000 tonnes/ year [1]. To establish a balance, recycling the waste composites, recovering the valuable carbon fibers (CFs), and reusing them into new composites is the only sustainable option [2].

Previously, landfilling and incineration were the popular disposal methods for their composites. However, various studies have proved that recycled carbon fibers (rCFs) to be close to their virgin properties, recycling industries have invested in a sustainable alternative to recycling CFRP composite wastes. Thermal recycling processes such as pyrolysis and solvolysis using supercritical/ subcritical water or mild solvents have proved to be highly efficient [1]. Recently, a novel thermal recycling process [2] capable of recycling CFRP composite wastes with maximum efficiency has resulted in clean recycled fibers without disturbing the fibers' structural integrity (fiber direction, arrangement, and length). However, all these processes exist on a laboratory scale. Among them, pyrolysis has been successfully established on an industrial scale in The UK (Gen 2 Carbon), Belgium (Procotex), and Germany (CFK valley & SGL group).

Fig. 1 presents the operating principle of the pyrolysis recycling process to recover CFRP composite wastes. First, the composite waste is size reduced using mechanical shredding and feeds into the system. The pyrolysis reactor is a closed chamber with no oxygen present. However, the process is done in an inert gas atmosphere to separate the valuable CFs from the matrix at 550 °C for the required time (depends on the waste quantity). Then, the recovered fibers are passed to a secondary heating chamber at 200 °C to oxidize the resin residues. Finally, the CFs are recycled, acquiring pyrolytic oil (can be used as feedstock) and hot gas (can be regenerated) [1]. Overall, industrial-scale pyrolysis possesses enormous sustainable benefits.

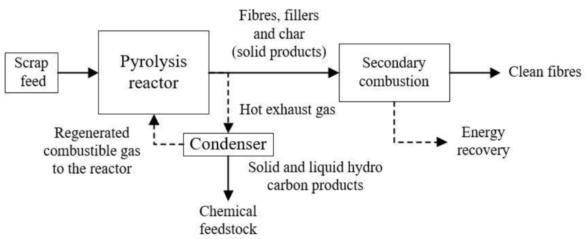



**Fig. 1.** Overview of the pyrolysis process [1]

The EU's regulations have moved from landfilling and incineration towards sustainable recycling regulations. In which recycling processes with low emissions in carbon footprints are preferred [1]. So, industries utilizing the pyrolysis process to recycle CFRP composite waste have to reconsider advanced emission monitoring along with their primary monitorable parameters such as pyrolysis temperature, pressure, inert gas flow, heat regeneration, recycled fiber quality, etc. Also, considering the type of CFRP composite wastes cumulating, which lacks composite information such as profile, the volume of fiber and resin, composite type, resin type, etc., there is a need to implement an advanced monitoring system to moderate such industrial-scale process.

These strategies call for more control and automation. Therefore, applying IoT scheme frameworks can improve the overall connectivity bringing significant control, monitoring, and safety to the process [3]. However, the existing semi-automated and, in some cases, fully automated pyrolysis plants need to be continuously updated and constantly monitored to couple with the varying CFRP composite waste types. To overcome such research gaps and eradicate manual errors, Internet-of-Things (IoT) based framework is proposed. Typically, an IoT is a network of devices interconnected to each other using some communication technology and using sensors and actuators to gather data from different devices and send that data to the cloud to store, process, and get information [4].

This paper studies the theoretical implementation of an Internet-of-Things (IoT)-based framework into the pyrolysis process to recycle CFRP composite wastes and manage the process based on the principles of cyber-physical systems. The proposed framework consists of sensors and actuators to collect the required data and communicate with central management constructed as an IoT platform that will articulate and manipulate data to satisfy the requirements of the recycling process, computationally modeled through logical relations between physical entities. Furthermore, this study will focus on selecting a suitable IoT platform based on the requirements of the pyrolysis process. Incorporating IoT and its platform into the pyrolysis process will improvize the plant by automatically calculating the mass of the CFRP composite wastes and adjusting parameters such as: operating time, temperature, atmospheric pressure, inert gas flow, increase energy efficiency, and reduce process flow time.

## 2 General IoT framework for recycling carbon fibers

The focus on implementing IoT technology into industrial-scale pyrolysis was discussed in the following section. First, an insight into improving pyrolysis factors such as cost, speed, reliability, scalability, and performance resulting from IoT was investigated [5]. Furthermore, discussions on IoT collected data as a reliable source of information in current decision-making and future improvements into the pyrolysis process.

### 2.1 Industrial-scale pyrolysis process and IoT framework

Fig. 2 presents the implemented IoT network in an industrial-scale pyrolysis setup. The process includes sensors and actuators at all the crucial sections (starting from the waste feed till the CF recovering) throughout the process. The IoT framework is designed to be flexible and extended to multiple furnaces (pyrolysis process1, pyrolysis process2…... pyrolysis process N). Fig. 1 presents the pyrolysis process setup. The gathered data from the individual furnaces are sent to the main gateway and forwarded to the cloud for storage, processing, and visualization. In addition, information can be taken from data stored in the cloud. Based on that data, a further future business decision can be taken that can improve and speed up the CF recycling process with precision and accuracy.

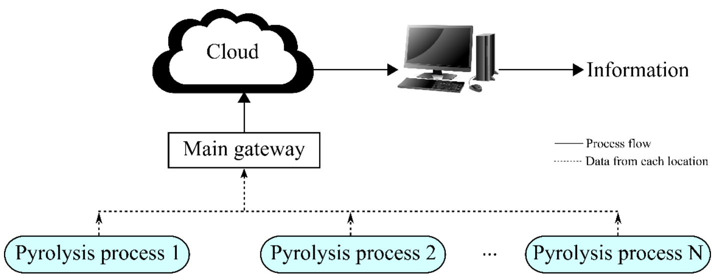



**Fig. 2.** IoT framework implementation in industrial-scale pyrolysis

### 2.2 The role of IoT in process monitoring

CFRP waste recycling process will be benefited from an IoT deployment by adding physical elements to the cyberworld on what is known as a Cyber-Physical System (CPS), as introduced in [4]. This will enable better control of the input material transported along with the conveyor, control temperature of the furnace, and real-time monitoring of char production. In Fig. 3, It can be seen how IoT is used to form a cyber-physical environment where sensing and actuators are connected to the gateway making the interface to the cloud server. In the cyber world, a manager can remotely perform monitoring, control, or analyze stored data.

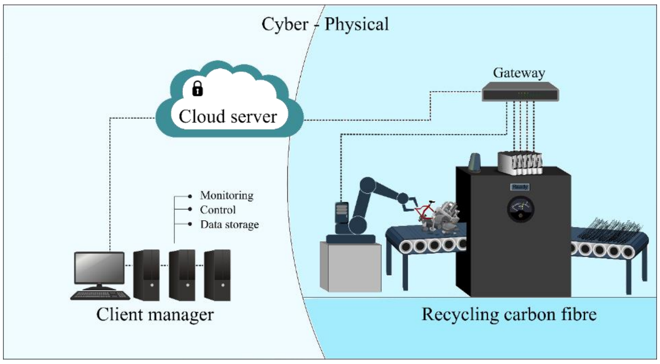



**Fig. 3.** IoT schematic of an in-site CF recycling process

### 2.3 Selection of IoT platform for pyrolysis process

Fig. 4 presents the framework of IoT platforms. IoT application needs a platform to run smoothly and provide the data to make future decisions based on the data received from the IoT platform. Hundreds of IoT platforms are available, and finding a suitable IoT platform for a specific IoT application is complex. A lack of experience and knowledge compounds the problem, and in some cases, a company may select a platform without adequate requirements analysis, which later leads to problems [6]. Companies can select an appropriate IoT platform for their IoT application if they first analyze their business requirements and start selecting the IoT platform with precise business requirements and knowing key factors of IoT platforms [5].

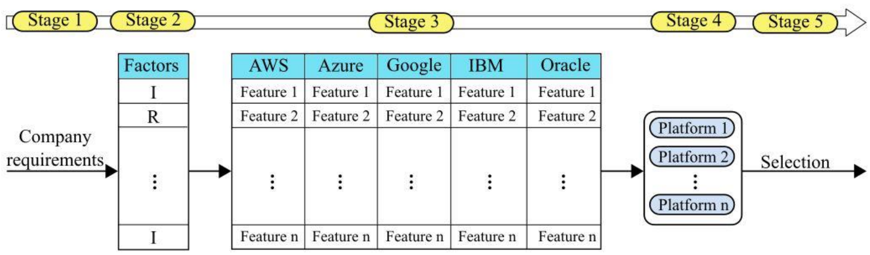



**Fig. 4.** Framework for the selection of the IoT platform

The process for the selection of an IoT platform has five stages. In stage 1, the company identify and finalizes their business requirements. In stage 2, the identified requirements are categorized as "Important" or "Required". In stage 3, the requirements are compared with the features provided by the IoT platforms. Only the five most important IoT platforms based on the market shares are selected in this case. In stage 4, IoT platforms are selected that are capable of fulfilling the requirements. Finally, in stage 5, a suitable platform is selected for the business IoT application.

The requirements for the pyrolysis process to recycle CFRP composite waste are to eliminate the manual work involved in the process, speed up the recycling time (> 1 tonnes/day), increase energy efficiency, and increase the overall recycling efficiency of the process. Furthermore, it calculates the amount of heat required for a specific amount of CFRP composite waste and estimates the amount of heat generated during the process. Finally, calculates the amount of emissions (exhaust and later outside the system). In short, the pyrolysis process needs stability, flexibility, scalability, security, attractive interface, data analytics, and interoperability throughout the system process. The process also needs and user-friendly application development environment for its IoT business application in the cloud.

The requirements of the pyrolysis process are categorized into "Important" and "Required" factors. The requirements stability, flexibility, scalability, security, attractive interface, and data analytics are considered "Important", and the requirements interoperability is considered "Required". The requirements of the pyrolysis process are compared with the features provided by the selected five major IoT platforms, as shown in Table 1. All the essential requirements of the pyrolysis process were matching with the features provided by the AWS IoT platform. However, Azure lacked the required flexibility and the required factor Interoperability.

**Table 1.** IoT platform features adopted from [5]

| Factors | AWS | Azure | Google Cloud | IBM Watson | Oracle IoT |
|---|---|---|---|---|---|
| Scalability | Yes | Yes | yes | yes | yes |
| Flexibility | Yes | - | yes | - | yes |
| Data analytics | Yes | yes | yes | yes | yes |
| Disaster recovery | Yes | yes | no | no | no |
| Stability | Yes | yes | yes | - | - |
| Security | high | high | high | high | high |
| Data ownership | - | yes | - | - | - |
| Protocol support | Yes | yes | - | yes | yes |
| System performance | Yes | - | yes | yes | - |
| Time to market | Yes | yes | - | - | yes |
| Legacy architecture | Yes | - | - | - | yes |
| Attractive interface | Yes | yes | - | no | - |
| Pricing model | bad | bad | good | - | - |
| Cloud ownership | Yes | yes | yes | - | yes |
| Interoperability | Yes | - | - | - | yes |
| App. environment | Yes | yes | yes | yes | yes |
| Hybrid cloud | Yes | yes | - | - | - |
| Platform migration | Yes | yes | - | - | - |
| Previous experience | Yes | yes | - | - | - |
| Edge intelligence | Yes | yes | yes | - | yes |
| Bandwidth | - | - | good | - | - |

### 2.4 Communication requirements

Currently, wireless communication advances have countless open opportunities for industrial applications. They can be critical enablers for monitoring and controlling impossible tasks before due to low flexibility and cost [7]. For the past decades, requirements for the industrial-scale process have been discussed, but they are bounded to every application. 5G cellular communication, evolution from previous 4G networks, has come with an improved set of characteristics that can increase the operational performance of industrial applications. Despite all the advantages, challenges such as interoperability, quality of service, ease of use, reach, cost, and security remain essential goals to be investigated to ensure overall benefit [8]. Please note that the first paragraph of a section or subsection is not indented.

5G technology achieves superiority due to its main three cornerstones: massive machine-type communications (mMTC), ultra-reliable and low latency communications (URLLC) Enhanced mobile broadband (eMBB) [3]. In Fig. 2, the communication means can be achieved by 5G wireless communication in the recycling process environment. The main argument for choosing this is scalability, flexibility, and cost. A wired communication system will need physical installation architecture to increase the cost of a component's change significantly. The communication requirements for the recycling process are as follow: (a) update time (process data) 1.5 ms, (b) transmission time (process data) 0.5 ms, (c) distance between logical endpoints 250 m, (d) reliability (redundancy), (e) less than 50 devices connected at the same time. The requirements can be achieved by installing one 5G base station inside the recycling building to ensure full connectivity between devices and cloud servers and allow remote operability.

## 3 Conclusion

The industrial-scale pyrolysis process to recycle CFRP composite wastes was studied by implementing IoT technology. The study proposes implementing sensors and actuators to collect data from the recycling plant and keeping in mind that the incoming composite wastes to be recycled will have varying composite types and properties. Therefore, the proposed IoT framework will adopt the recycling process conditions according to the composites to unify the standard pyrolysis setup. Furthermore, the methods involved in selecting a suitable IoT platform for the data collection, processing, storage, and visualization of results were also discussed. Overall, implementing the proposed IoT framework will enhance the recycling process adaptability over various CFRP composite waste types. As a result, the proposed IoT framework will facilitate the pyrolysis process into efficient energy utilization, reliability, scalability, and reduce the overall recycling time and cost. Furthermore, the information generated from the data collection and processing using the IoT platform can monitor and maintain the pyrolysis plant. Also, it is capable of influencing current and future recycling business decisions.

## References


1. Karuppannan Gopalraj, S., Kärki, T.: A review on the recycling of waste carbon fibre/glass fibre-reinforced composites: fibre recovery, properties and life-cycle analysis. SN Applied Sciences. 2, 433 (2020). https://doi.org/10.1007/s42452-020-2195-4
2. Gopalraj, S.K., Kärki, T.: A Study to Investigate the Mechanical Properties of Recycled Carbon Fibre/Glass Fibre-Reinforced Epoxy Composites Using a Novel Thermal Recycling Process. Processes 2020, Vol. 8, Page 954. 8, 954 (2020). https://doi.org/10.3390/PR8080954
3. Narayanan, A., Sena, A.S. De, Gutierrez-Rojas, D., Melgarejo, D.C., Hussain, H.M., Ullah, M., Bayhan, S., Nardelli, P.H.J.: Key Advances in Pervasive Edge Computing for Industrial Internet of Things in 5G and Beyond. IEEE Access. 8, 206734–206754 (2020). https://doi.org/10.1109/ACCESS.2020.3037717
4. Gutierrez-Rojas, D., Ullah, M., Christou, I.T., Almeida, G., Nardelli, P.H.J., Carrillo, D., Sant'Ana, J.M., Alves, H., Dzaferagic, M., Chiumento, A., Kalalas, C.: Three-layer Approach to Detect Anomalies in Industrial Environments based on Machine Learning. (2020)
5. Ullah, M., Nardelli, P.H.J., Wolff, A., Smolander, K.: Twenty-One Key Factors to Choose an IoT Platform: Theoretical Framework and Its Applications. IEEE Internet of Things Journal. 7, 10111–10119 (2020). https://doi.org/10.1109/JIOT.2020.3000056
6. Ullah, M., Smolander, K.: Highlighting the Key Factors of an IoT Platform. In: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). pp. 901–906. IEEE (2019)
7. Underberg, L., Willmann, S.: Categorization of industrial communication requirements as key to developing application profiles. IFAC-PapersOnLine. 53, 8297–8302 (2020). https://doi.org/10.1016/J.IFACOL.2020.12.1919
8. Khaled, Z. El, Mcheick, H.: Challenges and requirements for communication technologies in connected zones. Procedia Computer Science. 155, 27–34 (2019). https://doi.org/10.1016/J.PROCS.2019.08.008

# Publication V

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P. H. J

**Unified Framework to Select an IoT Platform for Industrial Energy Management Systems**

Reprinted with permission from
*IEEE MIPRO 44th International convention*
pp. 950–955, doi: 10.23919/MIPRO52101.2021.9597128

# Unified Framework to Select an IoT Platform for Industrial Energy Management Systems

Mehar Ullah, Arun Narayanan, Annika Wolff, Pedro H. J. Nardelli
LUT University, Lappeenranta, Finland
mehar.ullah@lut.fi

***Abstract*—The world population is increasing at a rapid speed and so are the requirements for everyday life like energy, technology, and industries. In the past few years, there has been a huge increase and development in the industrial sector and new technology like the Internet of Things (IoT) has played a major role in it. The industrial sector is consuming a large portion of the energy worldwide and is contributing a large amount of greenhouse gas emissions. Hence, they face significant economic, social and environmental pressures to create energy-efficient processes and systems of production and directly manage their energy consumption, looking at aspects beyond direct costs. One of the effective ways to reduce energy consumption, cost, and greenhouse gas emissions is to use an industrial energy management system (IEnMS). The IEnMS uses IoT for (big) data collection, processing, storing, and visualization. In this case, one key challenge the industries face is the selection of an IoT platform among the hundreds of IoT platforms in the market. In this paper, we adapt our previously proposed general framework to choose an IoT platform focusing here on the requirements of IEnMS. The proposed framework provides an objective methodology that can be used to select the most suitable IoT platform for different IEnMS based on their particular requirements.**

***Keywords*— *IoT, IoT platform, platform features, industrial energy management system, framework***

## I. Introduction

Energy and climate change are one of the most important topics in the public eye. Across every sector whether it is industry, housing or government, there is an attempt to save energy and control climate changes [1] because of the fact that energy is finite and can be used in a smart and efficient way to protect the current and future generations [2]. Industry is the one that is consuming the highest amount of energy and it is estimated that about 42.3% of the total energy produced in the world is consumed by the industrial sector [3]. For a company to be in the race of competition and position themselves as a market leader in the long run, the company has to give priority to the management and operation of the company's energy system. Due to this reason industry needs more attention to use effective energy management.

Energy management is one of the promising means for an industry to reduce energy consumption and related energy cost and reduce carbon emission. Back in nineties, energy cost was a small factor for the industrial companies, this is due to the reason that the energy cost was very low as compared to the production size and greenhouse gas emission was not the priority. But within the last decades the energy and energy sourcing prices in Europe increased rapidly. Industrial sector is focusing on energy management in which the company prepares long term energy policies, planning and targeting for the energy use. Industrial Energy Management System (IEnMS) is used to implement this whole process and control energy consumption, green house emission, reduction in energy use by a specific percentage and save cost.

The Internet of Things (IoT) is promising to make the environment smarter and connecting devices to each other and with the platform for communication and data exchange. Connected devices are adopting IoT at a rapid pace. Since IoT is expanding continuously, the need for connectivity methods grows to ensure that the devices work, capture, analyze and manage the data accurately. For efficient connections and sharing of information, communications protocols like MQTT and HTTP are used [4] . The best way to store, process and analyse that data within its infrastructure is by using an IoT platform that comes with the data processing capabilities [5]. It will not only allow you to collect data from Industrial machines but can also help you set up a custom visualization dashboard.

This contribution aims at utilizing our developed theoretical framework for selecting a suitable IoT platform for IEnMS by using IoT and big-data. Our goal is to provide an objective while general methodology that different industries can apply when selecting the most suitable IoT platform for IEnMS based on their particular needs and business requirements. In other words, this paper will support industries to carry out a detailed analysis of their own energy requirements and understand the key components of their Energy Management (EnM) to find the best match for their IEnMS. The study will find the answers of the following questions. 1) What are the components of EnM in industrial sector? 2) How are the Industrial energy management systems working using IoT and big-data? 3) How to design a framework for selecting a suitable IoT platform for industrial energy management system by comparing the requirements of IEnMS with the features provided by the IoT platforms?

The rest of the paper is organized as follows. Sec. II contains Industrial EnM (IEnM). Sec. III explains IEnMS and in same section we have explained Industrial Energy Management system its working. Sec. IV contains Internet of Things (IoT), IoT platform and the need for IoT platform. Sec. V contains our proposed theoretical framework for selecting IoT framework for IEnMS. Sec. VI concludes the paper.

## II. Industrial Energy Management (IEnM)

One of the indicators of the economic development is the fast growth in industrial sector. With the passage of time the number of industries is growing and so is the need of energy [6]. The increase in the use of energy will be continuous with the rapid development in the industrial sector and the fast growth of the world population and will create problems like increase in energy price and the carbon emission. Industry needs energy management to use energy efficiently and also control carbon emission. Energy management consists of practices and processes to improve energy efficiency. As a result, in the recent years many organizations have put focus on energy related issues to improve their productions and operations and improve energy efficiency and thereby reduce energy usage and energy cost. Industry needs energy management to use energy efficiently and also control carbon emission.

### A. Components of IEnM

In this section we will discuss the important components of IEnM that we have identified from the literature [6]–[10]. There are five main components in IEnM planning/strategy, operation/implementation, controlling, organization and culture.

**Planning/Strategy**: The first phase of EnM has three parts. The *first part* is the written long-term energy policy of the company. *Second part* is energy planning and target setting in which the industry is making the plans and setting future targets for the energy use. The target may be the lowering the green house gas emission, energy consumption, reduction in energy use by specific percentage etc.. *Third part* is the strategic energy risk management that is used to analyze any type of risk the company can face related to the energy use and how the risk can be managed by the company's predetermined financial objectives and risk tolerance.

**Operation/Implementation**: The second phase of EnM that consist of three parts. The *first* part is Implementation of energy efficiency measure in which the companies are implement specific energy projects and energy efficiency technologies to reduce electricity consumption. The *second part* is the investment decision on energy efficiency measure in which companies are conducting a systematic economic calculation to calculate the return of investment. The *third part* is the energy audit, in which the companies within the operations of an energy management constantly review the status-quo and highlight energy saving potential.

**Controlling**: The third phase of EnM that consist of three parts energy accounting, performance measurement and bench-marking. *Energy accounting* is the process of constant analysis of the energy consumption and measure the energy efficiency monitoring on a regular basis and is reported. *Performance measurement* is an integral part of EnM and it defines the key performance indicators (KPI) for energy efficiency. KPI's are describing the relationship between an activity and the required energy. *Energy benchmarking* is an activity that is focusing on energy performance and can be defined as the method used to compare the energy efficiency between or within entities and is a useful contribution to reduction in energy use and related cost and emission. There are three types of bench-marking industrial benchmark, historical benchmark and company-wise benchmark. In industrial benchmark, the company compares its own facilities and process with the facilities and process of other company. In historical benchmarking the company compares its energy consumption of a process or facility with its own process or facility in the early times. In company-wise bench-marking the company compares its facilities and processes inside the company.

**Organization**: The fourth phase of energy management that consist of two parts. *Selection of energy manager* which is based on the experience and should be climate friendly. There should be a close link between the energy manager and the top management, and is responsible to update the top management about the activities and progress of the energy management. The second part is the *integration and standardization*. According to industrial companies energy management should be integrated with the production management processes using ICT tools and standardization. The production process and evaluation of potential energy saving investment can be controlled by the ICT tools and the transparency of the industrial companies can be increased by the standardization.

**Culture**: The last phase of energy management consists of two parts. First part is *education and training*. In industries at the corporate level or plant level the energy manager needs man-power which require basic education to meet the requirements. For this purpose, the industrial companies need continuous energy related training, which provides positive impacts on the energy management. *Staff motivation* is the second part in which the industrial companies needs to motivate the staff to actively participate in improving the energy efficiency. In that case the companies give rewards to the technical and operational staff, which helps to sustain the momentum and improve the overall support for the energy management program.

## III. Fundamentals of IEnMS

For an organization to improve its energy performance, one of the best way is to use an Industrial Energy Management System (IEnMS). It monitors, controls, and optimizes energy performance in a plant and measures the consumption of energy. It is also used to diagnose problems like over-consumption and leaks across the entire plant. For energy consumers including industrial, public sector and commercial organizations an EnMS is a framework to manage their energy usage and can be defined as "a set of interrelated or interacting elements of a plan which sets an energy efficiency objective and a strategy to achieve that objective" [11]. it provides companies the opportunities to improve energy savings by adopting energy saving technologies. In most cases for a successful implementation of EnMS required low investment cost and specialized expertise and staff training.

TABLE I
REQUIREMENTS OF INDUSTRIAL ENERGY MANAGEMENT [7]

| S No | Requirements for IEnM |
|---|---|
| 1 | Development and implementation of strategic plan that includes the energy policy and specific targets for the energy savings. |
| 2 | Organizing different energy activities including the allocation of responsibilities and tasks. |
| 3 | Establishment of management team that is leading by an energy manager who will be responsible to report directly to the high management. |
| 4 | Development of related policies and procedures which can include energy procurement, energy usage, cheap energy purchases etc. |
| 5 | Carrying out the initial energy audit to identify energy saving potentials |
| 6 | Planning and implementation of energy efficient measures |
| 7 | Identification of the company's unique key performance indicators which can show the measure progress on a regular basis. |
| 8 | Implementation of the energy meters for the monitoring of the energy consumption at the main production processes at a regular interval. |
| 9 | Reporting of the information gathered from the data to the high management. |
| 10 | Providing progress to the high management so that the management shows interest in the energy management activities. |
| 11 | Training, motivating, and proving information to the employees of the company about energy management activities. |

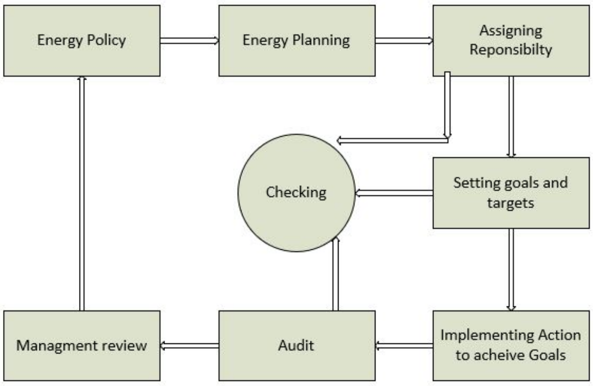



Fig. 1. Industrial Energy Management

Industries are required to follow a series of defined steps to establish a successful EnM. The main steps required are development of energy policy and assigning responsibilities, highlighting main energy users, setting measurable goals and targets, implementing actions to achieve the goals, checking actions are successful, and continuous system review as shown in Figure 1.

### A. Working of Industrial Energy Management System

An Industrial Energy Management System is a computer-based system that is used to collect and measure energy data from the field like HVAC units (heating, ventilation and air-conditioning), lighting system, water and gas meters installed on production line etc. and provide that information to the user. The working of IEnMS is explained below.

Building data collection strategy: A system that is used to collect real time accurate and granular data and have the information that where and when the energy has been used and by which device(machine). The data is collected by installing sensors/smart meters, and sub-meters on the incoming supply and the large energy consumer (device). For example, on the HVAC, production line, boilers etc. The objective of this part is to monitor the collection of real-time data that shows where the majority of the energy being used.

Transform raw energy data into useful information: In this phase the collected data is analyzed, interpreted and is converted into useful information. Here the big-data software is used that can easily import the raw data from different machines using IoT devices and then convert the raw data into useful information in the forms of charts, graphs etc. that is user friendly. Here the raw data collected may related to the production levels, weather data, humanity and other factors that could influence energy usage to generate the key performance indicators.

Assign responsibility, analyze data: In this phase the information provided needs to be converted into useful and meaningful reports, that can only be possible by adding the information to the knowledge of the facility and some expertise in energy management. This can be done by the energy manager. The role of the energy manager is to interpret the information provided by the Industrial Energy Management System and combine this information with the company's process, set the targets.

Interpret the results, and agree an action plan: Here in this phase the energy usage reports are available to the energy manager. The energy manager and energy management team start communication with the departments to start an energy policy and make agree to an action plan.

## IV. INTERNET OF THINGS AND IOT PLATFORM

The term "Internet of Things", or "IoT", was introduced by the British technology entrepreneur Kevin Ashton in 1999 as the title of a presentation at Procter and Gamble.

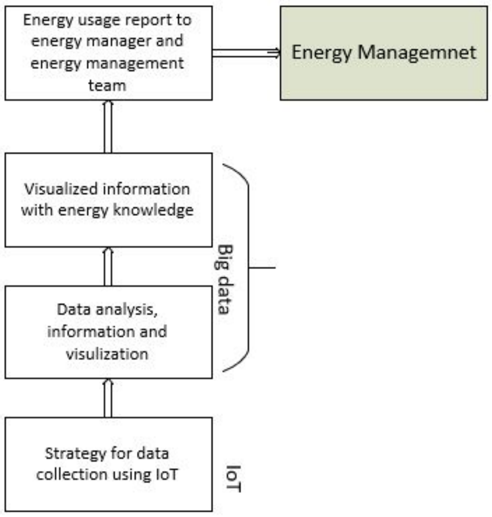



Fig. 2. Industrial Energy Management System

IoT and can be defined as small and complex systems that allow businesses, governments, and citizens to adopt and interconnect physical objects and virtual objects based on existing and evolving interoperable information and communication technologies [5]. IoT is a new technology paradigm that has emerged as a global network of machines and devices capable of interacting with each other and with the platform for collecting, analyzing, storing and visualizing data generated from the devices and machines using sensors, actuators, communications, and analytical tools [12]. IoT is now playing a main role in Industrial sector by collecting data from machines and other sources [13] and IoT platform is used to store, process and perform analytics on the collected data and provide useful information to be sent to the customer, as shown in Figure 3.

An IoT platform provides services to IoT devices and customers and enables IoT device an endpoint management, connectivity and network management, processing and analysis, data management, application development, access control, security, monitoring, event processing and interfacing [5]. Recent increase and development in mobile devices, embedded technologies, cloud computing and data analytics has resulted in a boom in IoT utilization, in terms of personal and organizational use, to conduct information exchange to facilitate recognition, monitoring, tracing, positioning and administration [5],[6]. The number of IoT platforms are increasing at a rapid speed, For example, in 2015 the number of IoT platforms were almost 260, which grew to 360 in 2016, exceeded to 450 in 2017 and reached to 620 in 2019 [14]. Requirements for IoT platforms, which provide important services and features for IoT applications, change as new IoT devices emerge [9]. This complexity in the context of rapid change poses

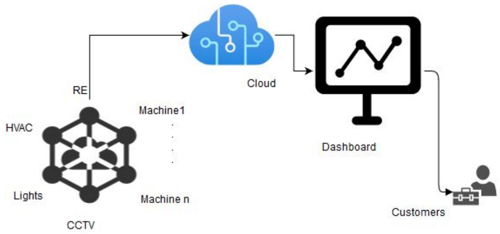



Fig. 3. IoT for IEnMS

challenges for businesses, governments and citizens that often have little experience of the infrastructure of IoT and limited knowledge of how to select an IoT platform that can meet their current and future needs.

IoT application needs a platform to run smoothly and provide the data so that the companies take present and future business decisions based on the data received from the IoT platform [11]. An IoT platform is the main part of an IoT solutions. Among the hundreds of IoT platform vendors in the market, it becomes very difficult for the companies to find and select a suitable IoT platform that is best suitable for their business application and which can fulfil their maximum business requirements. This problem can be solved by two steps. In the first step, the company can identify their complete present and future business requirements. In the second step the company should have the knowledge of some of the key factors of an IoT platform before selection. In this way the company can select a suitable IoT platform for their business application by comparing the requirements with their features.

## V. General framework for selection of IoT platform for IEnMS

We have aimed to create a more general approach that can be more widely used across all the cases. To show how our general framework can be applied to assessing and choosing an IoT platforms, in this study we have selected the top five IoT platforms (AWS, Azure, Google, IBM, and Oracle) based on market share. We have compared these IoT platforms according to the twenty-one key IoT platform factors that we have identified in [5]. We have compared these twenty-one key factors with the features provided by those selected five IoT platforms as shown in Table II.

More specifically, the entries of Table II have the following meaning related to the specific feature to be considered: 'yes' means the feature is available, 'high' indicates strong, 'bad' shows weak, 'good' indicates that the feature is very good, '-' shows that the feature is unknown and 'no' indicates that the feature is not available in the platform. In order to identify and fill the features of the selected five IoT platforms, different articles [5], [15]–[20] have been studied from many databases. Some websites [21]–[26] have been used, especially the websites of those selected IoT platforms. A few white papers [27] have also been studied.

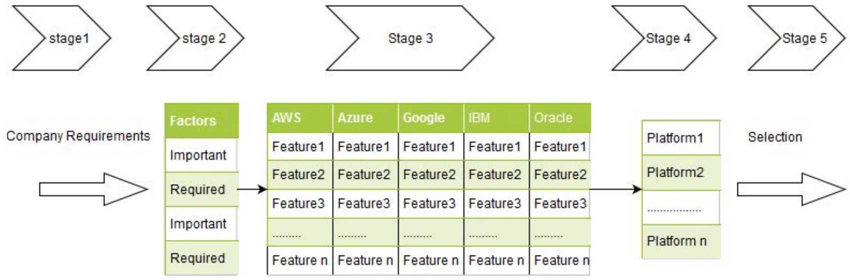



Fig. 4. IoT platform selection Framework for Industrial Energy Management System

TABLE II
FEATURES PROVIDED BY IOT PLATFORMS [13]

| Factors | AWS | Azure | Google | IBM | Oracle |
|---|---|---|---|---|---|
| Scalability | Yes | Yes | Yes | Yes | Yes |
| Flexibility | Yes | - | Yes | - | Yes |
| Stability | Yes | Yes | Yes | - | - |
| Security | High | High | High | High | High |
| Data analytics | Yes | Yes | Yes | Yes | Yes |
| Disaster recovery | Yes | Yes | No | No | No |
| Data ownership | - | Yes | - | - | - |
| Protocol support | Yes | Yes | - | Yes | Yes |
| System performance | Yes | - | Yes | Yes | - |
| Interoperability | Yes | - | - | - | Yes |
| App. environment | Yes | Yes | Yes | Yes | Yes |
| Cloud ownership | Yes | Yes | Yes | | Yes |
| Pricing model | Bad | Bad | Good | - | - |
| Legacy architecture | Yes | - | - | - | Yes |
| Attractive interface | Yes | Yes | - | No | - |
| Time to market | Yes | Yes | | | Yes |
| Bandwidth | - | - | Good | - | - |
| Edge intelligence | Yes | Yes | Yes | - | Yes |
| Hybrid cloud | Yes | Yes | - | - | - |
| Platform migration | Yes | Yes | - | - | - |
| Previous experience | Yes | Yes | - | - | - |

The framework for selection of an IoT platform is illustrated in Fig.4. as a schematic of the selection procedure. The whole process consists of five stages. In the first stage, the company finalize their business requirements. In the second stage, the company requirements are applied to prioritising which factors are required (R) and important (I) for this business context. In the third stage, the R and I factors are compared with the features provided by the five selected IoT platforms. The IoT platform/s that provide a maximum of the features as compared to the requirements are selected and shifted to the stage four. In stage four, there might be one or many IoT platforms that match the required and important factors. Stage five is the decision, which is explained next.

If there is one IoT platform that provides the most required and important features then the same IoT platform can be selected for the business application. But, if there are multiple IoT platforms providing these features then the company may choose an IoT platform based on the comparison of their match to "I" factors and select a suitable IoT platform for their business needs. There might also be chances that none of the platforms provide all of the required features; this might indicate that new platforms should be selected and evaluated accordingly.

For the IEnMS, the five stages of the framework are explained in a simple way. A company is interested to start energy management system to save reduce greenhouse gas emission, save energy and reduce energy bills. The company is interested to use IoT application for its EnMS. Initially the have to learn the components of IoT to understand what is IoT and how it works. Secondly they have to study to know what is an IoT platform and what are key factors of an IoT platform. They need an IoT platform for their business application but they do not know which platforms are providing what features and which one is best. when the company have the knowledge of IoT, IoT platform factors and the features those platforms are providing then in stage 1, the company go through each of the 21 factors that have been identified as important in choosing a platform and use this to help them to formulate their business requirements. In stage 2, these factors are prioritized as being either required (R) and important (I) for their business needs. They find that their required factors to consider (R) are for example, scalability, stability, system performance, attractive interface, edge intelligence, time to market, flexibility, and previous experience. Their important factors to consider (I) are for example pricing, security, data analytics, disaster recovery, and interoperability. In stage 3, the R and I factors are compared with identified features of IoT platforms. The platform that is fulfilling the "I" and "R" requirements of the company can be selected. In case there are multiple of IoT platforms that are offering completely the "I" and "R" requirements of the company, then the company can select one that is providing the "I" requirements in a better way. In some cases it can happen that non of the IoT platform are providing all the "I" and "R" requirements, in this case the company can search some other IoT platform for their IoT application.

## VI. CONCLUSIONS

The aim of this study is to build an objective methodology that can support industries to select the most suitable IoT platform for their IEnMS based on their specific needs. To do so, we first highlight the components of EnM and later we have explained the components and working of IEnMS. Second, we identified twenty-one key factors of IoT platforms from the literature. Finally, we have designed a theoretical framework for selection of IoT platform for IEnMS and tested it in five well-known examples. This research then provides a general framework to select the most suitable IoT platform for industries to build their IEnMS by comparing its specific requirements with the features offered by the different platforms.

We believe industries can select an appropriate IoT platform for their IoT application, if they first analyze their business requirements, and start selection of the IoT platform with clear business requirements and have the knowledge of key factors of IoT platforms. This study highlighted important key factors of IoT platform. Those key factors may cover some of the current requirements of the IEnMS and can help ensure that current and future needs of the business are meet. It will facilitate the companies to select a suitable IoT platform for their business needs.

## ACKNOWLEDGEMENTS

This paper is supported by Academy of Finland via: (a) ee-IoT project n.319009, (b) FIREMAN consortium CHIST-ERA/n.326270, and (c) EnergyNet Research Fellowship n.321265/n.328869.

## REFERENCES

[1] E. Pinero, "Iso 50001: setting the standard for industrial energy management," *Green Manufacturing News*, pp. 21–24, 2009.

[2] A. Sa, P. Thollander, and E. Cagno, "Assessing the driving factors for energy management program adoption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 538–547, 2017.

[3] M. Wei, S. H. Hong, and M. Alam, "An iot-based energy-management platform for industrial facilities," *Applied energy*, vol. 164, pp. 607–619, 2016.

[4] M. Ullah, S. R. U. Kakakhel, T. Westerlund, A. Wolff, D. Carrillo, J. Plosila, and P. H. Nardelli, "Iot protocol selection for smart grid applications: Merging qualitative and quantitative metrics," in *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2020, pp. 993–998.

[5] M. Ullah and K. Smolander, "Highlighting the Key Factors of an IoT Platform," *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 901–906, 2019.

[6] B. Purwanggono, K. Ferastra, and A. Bachtiar, "Critical success factors evaluation of the iso 50001 energy management system implementation (case study: Pt. apac inti corpora, bawen, semarang indonesia)," in *IOP Conference Series: Materials Science and Engineering*, vol. 598, no. 1. IOP Publishing, 2019, p. 012115.

[7] M. Schulze, H. Nehler, M. Ottosson, and P. Thollander, "Energy management in industry–a systematic review of previous findings and an integrative conceptual framework," *Journal of cleaner production*, vol. 112, pp. 3692–3708, 2016.

[12] M. Ullah, A. Hekmatmanesh, D. Savchenko, R. Moioli, P. Nardelli, H. Handroos, and H. Wu, "Providing facilities in health care via brain-computer interface and internet of things," in *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2020, pp. 971–976.

[8] L. Sivill, J. Manninen, I. Hippinen, and P. Ahtila, "Success factors of energy management in energy-intensive industries: Development priority of energy performance measurement," *International journal of energy research*, vol. 37, no. 8, pp. 936–951, 2013.

[9] E. V. Tachmitzaki, E. A. Didaskalou, and D. A. Georgakellos, "Energy management practices' determinants in greek enterprises," *Sustainability*, vol. 12, no. 1, p. 133, 2020.

[10] A. Sa, P. Thollander, and M. Rafiee, "Industrial energy management systems and energy-related decision-making," *Energies*, vol. 11, no. 10, p. 2784, 2018.

[11] V. A. da Silva Gonçalves and F. J. M.-H. dos Santos, "Energy management system iso 50001: 2011 and energy management for sustainable development," *Energy Policy*, vol. 133, p. 110868, 2019.

[13] M. Ullah, P. H. Nardelli, A. Wolff, and K. Smolander, "Twenty-one key factors to choose an iot platform: Theoretical framework and its applications," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10 111–10 119, 2020.

[14] S. Liu, "• Number of IoT platforms worldwide 2015-2019 — Statista." [Online]. Available: https://www.statista.com/statistics/1101483/global-number-iot-platform/

[15] P. P. Ray, "A survey of IoT cloud platforms," *Future Computing and Informatics Journal*, vol. 1, no. 1-2, pp. 35–46, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.fcij.2017.02.001

[16] P. Agarwal and M. Alam, "Investigating IoT Middleware Platforms for Smart Application Development," pp. 1–14, 2018. [Online]. Available: http://arxiv.org/abs/1810.12292

[17] H. Hejazi, H. Rajab, T. Cinkler, and L. Lengyel, "Survey of platforms for massive IoT," *2018 IEEE International Conference on Future IoT Technologies, Future IoT 2018*, vol. 2018-Janua, pp. 1–8, 2018.

[18] M. Ammar and B. Crispo, "slimIoT : Scalable Lightweight Attestation Protocol For the Internet of Things."

[19] S. Challita, F. Zalila, C. Gourdin, and P. Merle, "A precise model for Google cloud platform," *Proceedings - 2018 IEEE International Conference on Cloud Engineering, IC2E 2018*, pp. 177–183, 2018.

[20] B. Nakhuva and T. Champaneria, "Study of Various Internet of Things Platforms," *International Journal of Computer Science & Engineering Survey*, vol. 6, no. 6, pp. 61–74, 2015.

[21] Amazon Web Services, "IoT Applications & Solutions — What is the Internet of Things (IoT)? — AWS," 2019. [Online]. Available: https://aws.amazon.com/iot/

[22] Google Cloud, "Google Cloud IoT - Fully managed IoT services — Google Cloud," 2019. [Online]. Available: https://cloud.google.com/solutions/iot/

[23] IBM Watson IoT Platform, "IBM Watson IoT Platform - Overview - Finland," 2019. [Online]. Available: https://www.ibm.com/fi-en/marketplace/internet-of-things-cloud

[24] Microsoft Azure IoT, "Microsoft Azure," 2019. [Online]. Available: https://azure.microsoft.com/en-in/pricing/details/iot-hub/

[25] Internet of Things, "Internet of Things (IoT) — Oracle," 2019. [Online]. Available: https://www.oracle.com/internet-of-things/

[26] IOTIFY, "Top 10 criteria to choose the best IoT cloud platform." [Online]. Available: https://iotify.io/top-10-selection-criteria-for-your-iot-cloud-platform/

[27] G. Cloud, I. Security, D. Overview, G. Cloud, G. Cloud, A. Layer, and T. Security, "Google Cloud Security Whitepapers," no. March, pp. 1–97, 2018.

# Publication VI

Ullah, M., Narayanan, A., Wolff, A., and Nardelli, P.

**Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing**

Reprinted with permission from
*IEEE MIPRO 44th International convention*
pp. 956–961, doi: 10.23919/MIPRO52101.2021.9596885

# Smart Grid Information Processes Using IoT and Big Data with Cloud and Edge Computing

Mehar Ullah*, Arun Narayanan*, Annika Wolff*, Pedro Nardelli*
* LUT University, Lappeenranta, Finland
mehar.ullah@lut.fi

***Abstract*—Smart grid applications typically use cloud computing to address the computational requirements for efficient electricity delivery. Recently, the emerging Internet of Things (IoT) has resulted in increased number of devices connected to the smart grid, including consumer gadgets, measurement equipment, and electrical and electronic devices such as smart power converters, phasor measurement units, and smart meters. These heterogeneous devices that are present in all the four stages of a smart grid—generation, transmission, distribution, and consumption—generate huge amount of structured, semi structured and unstructured data. Gathering, storing, and processing such huge data volumes using cloud computing creates problems of bandwidth, latency, disaster recovery, and cost. To overcome these problems, we present a theoretical discussion on the usage of IoT, edge computing and big data to collect the data from smart grid, process the data using edge computing and big-data analytics, and use the data for smart grid maintenance, energy information and future decisions. Further, we highlight how edge computing and big-data approaches have mitigated the above mentioned problems by shifting the control, intelligence, and trust to the edge of the network.**

***Keywords*— IoT, Smart grid, big data, edge computing, cloud computing***

## I. Introduction

The renewable energy-based smart grid (SG) concept emerged in the early 21st century, motivated by the necessity to increase renewable energy production in the electricity grid. The transformation of the 20th century traditional electric grid to modern 21st century smart electric grids was enabled by remarkable advancements in power electronics and internet and communication technologies (ICT). The integration of the new technologies enables efficient utilization of the energy production and consumption, providing opportunities for new energy resources like wind and solar etc., allowing the exchange of generated power from different sources and also bi-directional flow of power and communication. As a result, both utility companies and customers have been installing renewable energy sources (RES), such as solar and wind energies, inside the distribution grid [1]. Further, new technological paradigms such as Internet of Things (IoT) are influencing modern SG operations by improving communication, achieving better customer relationships, and handling the huge amount of data generated from the smart devices. In this sense, IoT is increasingly being used in SG applications for data gathering, communication, and smart analytics. For example, energy-based data analytics from the user to the utility can highly improve efficiency, reduce congestion, and improve power-supply reliability in 100% renewable-energy-based SGs in the future [1].

Recently, many industries across sectors have shifted their business operations to some form of cloud computing since cloud computing provides some benefits such as flexibility, operability, and cost savings. As a result, enormous amounts of data generated by devices are being sent to the cloud servers for processing and analysis [2]. In the industrial environment, such an increase in data from devices create many problems. First, a large amount of data is transferred to the cloud for analysis, but most of it may be irrelevant to the operations. Thus, this data transmission ends up creating high traffic to a central repository and increases costs due to extra unnecessary storage. Second, important data that need to be sent with small latency's of between seconds and millisecond can be very important for crucial industrial operations may experience costly delays. Third, sending data to cloud and retrieving that data can be very costly [3].

Edge computing is used to overcome the aforementioned problems in cloud computing [4]. The benefits of edge computing are that it moves data analysis and services away from centralized servers and a lot of data analysis is performed at the source of data collection [3]. Edge computing analyze data on the spot and filters the important data in real time; this improves the speed of data analysis and the decision-making process [5]. Edge computing is proving huge benefits to (IoT)-enabled business, but nevertheless, cloud computing remains important because having a centralized location for the data storage and analysis still has many benefits. In particular, non-time-sensitive data can be sent to the cloud, for example, for deep analysis post-hoc using machine learning (ML) methods to improve industrial operations and strategies [6]. Table I shows some differences between cloud and edge computing [6], [7].

The comprehensive sensing and processing abilities of IoT support many technologies in SG. Further, the rapid increase in IoT-enabled devices can cause explosive growth in data generation, resulting in the so-called "big data" regime, where the system generates data that is so large, fast, or complex that non-traditional methods are required for processing it. The generation of big data in SGs makes the existing data-processing capacities ineffective as edge computing does not have all the resources sufficient for the complex and intelligent big-data analytics tasks [8]. Hence, it is important to include more modern

TABLE I
DIFFERENCE BETWEEN CLOUD COMPUTING AND EDGE COMPUTING.

| Point of Difference | Cloud Computing | Edge Computing |
|---|---|---|
| **Operations** | Happens on the cloud platforms such as AWS, Azure, Google | Happens on the device itself or at the gateways |
| **Benefits** | Can store massive amount of data on scalable hosting on the cloud which can accessable anytime on the Internet | Network can be scalable independently with each new device that is added to the system, possibly working as a federation |
| **Suitable use case** | Suitable for the operations with more tolerance in terms of latency and requires high levels of computing power | This is suitable for low latency applications and that allows for distributed data storage, leading to a scalable and cost effective hosting providers |

big-data analytics to improve the data-processing capacity of IoT data [9]. Big-data analytics can be defined as the process by which the variety of IoT data are analyzed to find the trends, hidden patterns, unseen correlations and new information. This huge amount of data analysis and information gathering will provide benefits to companies for current and future effective decisions and will also provide benefits to the individual users [10].

In this paper, we present a theoretical framework using a combination of IoT, edge computing, big data, and analytics for the efficient collection of data from the huge number of devices connected to SG as well as the processing, storage, and visualization of the collected data. We elaborate on some benefits of edge computing and and big data, such as latency, bandwidth, disaster recovery, and price in the entire SG system, starting from data collection to data visualization. The rest of this paper is organized as follows. In Section II, we describe related studies on big-data analytics in smart grids and explain the concepts of IoT, edge computing, and big data. In Section III, we explain how these technologies—IoT, edge computing, big data, and big-data analytics—work together. Finally, Section IV concludes the paper.

## II. RELATED RESEARCH

In this section, we use systematic literature review structure, similar to [11], to answer the question "what type of technologies can be used to handle the massive amount of data from SG and obtain meaningful information from it, so that it can be used for better business decisions?" The main objectives are to identify technologies that can work together to effectively extract information from big data to fulfil the requirements of SG business and operations.

### *A. Smart Grid*

The SG technology concept has emerged to improve the flexibility and efficiency of the traditional grid and provide new opportunities for new generation methods such as wind, solar, and other RES based generation. The SG is essentially an electrical network that consists of infrastructure, software, and hardware, which enables it for two-way communication between all parts of the system and participants and efficiently generate power and enable distribution in the supply chain. As a self-sufficient distributed system that can provide energy from different sources including renewable and storage, SG also enables the suppliers and consumers access to the control and management capabilities [12].

SG allows two-way communication between the generation and consumption side with the help of devices such as smart meters, smart appliances, battery energy storage systems, power electronics converters, and other energy efficient resources. SG uses computer technologies for the improvements in automation, communications such as information exchange between consumers, transformers, and generation plants, and connectivity between many components of the power network, e.g., power gathered from different generation plants [13].

The SG works differently than the traditional grid; the network structure of the smart grid is complex having two-way communication and two-way interactions between the devices and the participant in the supply chain. The operations involves many steps from power generation to consumption, as explained below [14].

**Generation**: Power is generated from distributed sources that can include traditional power plants and renewable sources such as solar and wind. Electric storage can be used for generation-side management, including consumption-integrated storage such as electric vehicles.

**Transmission**: The generated power is transmitted using a network of transmission lines substations, and distribution systems. In the SG, transmission comprises three interactive components—smart control centers, smart power transmission networks, and smart substations. The smart transmission networks are conceptually built on the existing electric transmission infrastructure and the current advanced technologies—sensing, computing, communication and signal processing—provide services such as power utilization, power quality, network security, and reliability.

**Distribution**: The generated power is transmitted using a network of transmission lines that connects via substations to distribution systems that cover smaller areas and deliver power directly to the consumer. Because of the presence of both centralized and distributed power generators in modern SGs, the distribution networks have two-way electricity transmission, or, in other words, bidirectional power flow.

**Consumption**: The power consumption in modern SGs is often controllable and manageable by the end user using smart meters, sensors in appliances, plugs, and smart sockets. The user can control and manage their electricity consumption by using mobile phone applications or website applications to monitor and control the power usage.

**Control and management**: SGs have the capability

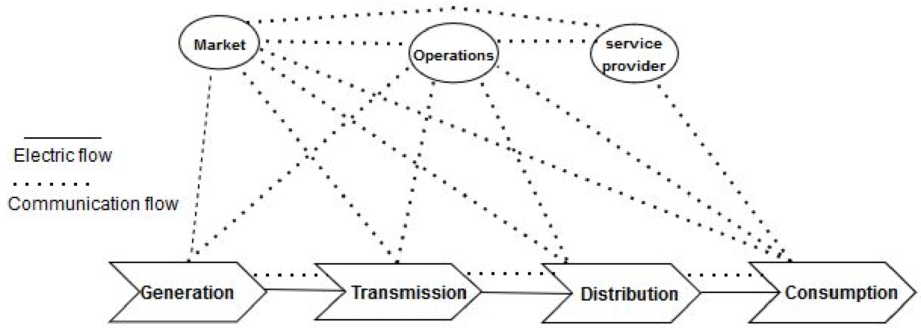



Fig. 1. Structure of Smart Grid (modified from [14]).

of control and management, and consumers, utility companies, and others in the energy industry can have a strong control on the energy usage and management. Data about the consumption and loads is generated from the connected homes, smart cities etc., and the information generated from that data is used by the companies and customers for their current and future decisions (using data analytics and visualization tools). For example, the energy companies can use the information for predictive maintenance; utility companies can use the information for demand and response programs; and residential users can use the information to reduce the energy consumption at the peak loads and reduce energy bills.

**Storage**: Electric storage is an important SG technology that enables generation-side management—households can store either extra produced energy or cheaper priced electricity, and later use it in the case of outage or when electricity is more expensive. Independent residential grids that are totally dependent on the renewable energy and generate a surplus of energy can store the surplus energy for future use.

### *B. Internet of Things*

The term "Internet of Things", or "IoT", was introduced by the British technology entrepreneur Kevin Ashton in 1999 as the title of a presentation at Procter and Gamble. IoT and can be defined as small and complex systems that allow businesses, governments, and citizens to adopt and interconnect physical objects and virtual objects based on existing and evolving interoperable information and communication technologies [15]. IoT is a new technology paradigm that has emerged as a global network of machines and devices capable of interacting with each other and with the platform for collecting, analyzing, storing and visualizing data generated from the devices and machines using sensors, actuators, communications, and analytical tools. IoT is now playing a main role in SG by collecting data from all the main phases of SG, including generation, transmission, distribution, and consumption [16], [17].

### *C. Edge computing*

Edge computing refers to the enabling technologies that allow computations to be performed at the edge of the network; from the cloud viewpoint, edge data is downstream data, and from the IoT services viewpoint, it is upstream data. In such a scenario, edge can be defined as any computing and networking resource that is between the data source and cloud. Smart phone is a simple example of an edge device, because it lies between the human user and the cloud. The aim of edge computing is to ensure that computations are performed at the proximity of data sources [2], [18].

The nodes at the network edge are performing many tasks such as data processing, caching, device management, and privacy protection to reduce the traffic from the devices to the cloud. In order to perform all these tasks in the network, the edge should be well designed to effectively meet the security, reliability and privacy protection requirements.

### *D. Big Data*

In simple terms, big data can be defined as the collection of unstructured, structured, and semi-structured data generated by the social media, devices, sensors, software applications, and digital devices that are continuously generating data [19]. The data collected is so large that the normal conventional data processing software and techniques are not able to process it. Big data is characterized by the three main determinants, called as 3Vs of big data—volume, velocity, and variety. Volume is the huge amount of data generated that make the datasets too large for the normal database technology. This type of data is measured in larger units of data, such as terabytes, petabytes, and exabytes. Velocity is the speed with which the data is generated, processed, and moved around in real time. Variety is the type of data (nature of data), i.e., whether the data is structured or unstructured [20].

The main idea of IoT is to connect heterogeneous objects to the internet and collect data from these devices, analyze the collected data, and make future decisions. Recently, due to dramatic improvements in the technology and business digitization, the number of devices connected

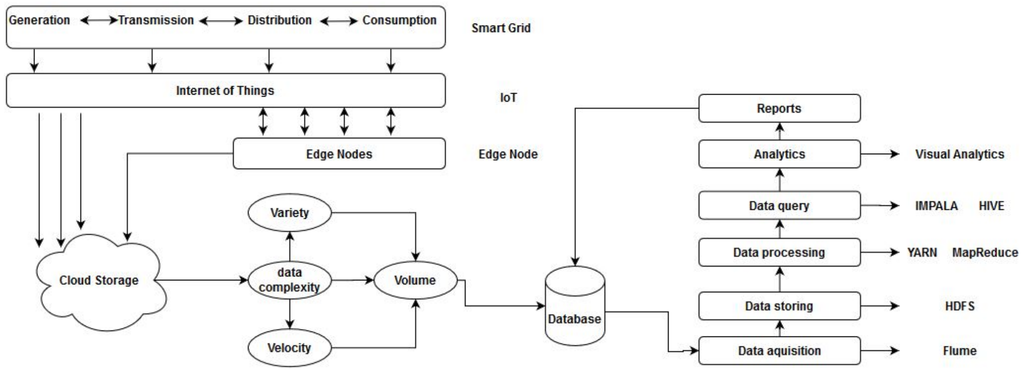



Fig. 2. IoT and Edge computing in Smart grids using Big data Analytic.

to IoT has increased tremendously; as a result, the amount of data has also increased tremendously so that there is a need to apply big data and big-data analytics to IoT. Big data and big-data analytics have high potential to extract meaningful information from the huge amount of data and improve the decision processes. The main requirements (functional and non-functional) of big data and analytics in IoT are explained below.

**Connectivity**: Connectivity in IoT is mostly ubiquitous with the heterogeneous objects in the network. Many objects are connected to internet via sensors in a smart environment. IoT services are mostly based on machine-to-machine (M2M) communication protocols that are required to handle a large number of streams, and it takes benefits directly from the cloud distributed storage and computing infrastructure [21]. The first and most important requirement of IoT is to provide reliable connectivity for big data and analytics. Reliable connectivity will provide big data and analytics the opportunity to efficiently combine and integrate the massive amounts of machine generated sensor data. Using the advanced wireless networks such as Wi-Fi and 4G/5G, many objects around us are able to connect to the computing and high performance infrastructure and facilitate the IoT services [22].

**Storage**: The amount of storage required for huge amounts of heterogeneous data in a low-cost hardware on a real-time basis has increased tremendously. The requirements of big-data storage in IoT are to handle massive amount of unstructured data and provide low latency for analytics. A challenge is that many sources of IoT data exist, for example, sensors' data, social media, etc., and they are modeled in various ways using different communication protocols and interfaces. Big-data technology provides some IoT-efficient data storage capabilities, but more robust solutions are required.

**Quality of services**: The ability to provide guarantee of a specific level of performance to the data flow is called Quality of service (QoS). The QoS provided by the IoT is that the IoT network should be reliable and should provide the guarantee of an efficient transfer of data from the sources that generates the big data. The QoS in the IoT network is very important to big data and analytics [23].

**Real time analytics**: IoT is growing rapidly and taking key steps to improve streaming analytics and provide timely decision processes. Real-time information about the IoT-connected objects are communicated and need to be analyzed in real time. Big data uses an operational database for the streaming data, and for most of the streaming data from web-enabled objects, big-data analytics performs real-time queries to extract information quickly, make decisions, and interact with the devices and people in real time [24].

**Benchmark**: Due to the fast digitization of businesses, many organizations have started to shift their business online using IoT. Many organizations are now facing challenges in storing and analyzing the huge amount of data connected through the IoT devices. Finding solutions to those challenges requires some deep understanding of the problems. Benchmark plays an important role in this situation by allowing the organizations to compare the quality of the big data and analytics solutions [25].

## III. THE ROLE OF IOT, EDGE COMPUTING,AND BIG DATA IN SMART GRIDS

To spur growth in businesses, effective business decisions are very important, and they are often made possible by getting information from collected data. IoT is a major source of data; by some estimates, there are currently more than ten billion devices connected to IoT networks[1], generating around trillions GB of data. These devices gather, analyze, share, and transmit data in real time. To handle such massive amounts of data, IoT needs edge computing and big data, making them the key to improve decision making [26].

In our this work, we have designed an architecture for getting the smart grid data using IoT and edge nodes as shown in Figure 2. The life-cycle of SG data starting from data generation to data analytics. The data is generated

[1]https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. Last access March 1, 2021.

from numerous smart meters, sensors, and digital devices with a specific time scale. The generated data may be from generation plants (wind farms, solar panels, conventional power plants, etc.), transmission and distribution networks (phasor measurement units, etc.), or customers (residential homes, electric vehicles, commercial buildings, factories, etc.). Data, such as weather, humidity, temperature, and pressure data, can also be collected from the environment. Some usable data, e.g., information about external events, can be collected from social media. Data generated from many sources increase the grid reliability. The generated data are transmitted to the IoT network using IoT devices such as sensors and actuators through the network technologies 3G/4G/5G, ZigBee, wi-fi, bluetooth, and wired communication.

The important data that needs to be processed quickly (requiring low latency) is processed by the edge nodes. The edge nodes are close to the data collection points, and therefore, require very low latency [3]. However, there are some cases in which the benefits may not be achieved, since the latency not only depends on the distance between the data collection point and edge processing server, but also on the edge server's processing power, tasks' computational complexity, and edge traffic [4]. Figure 3 shows the latency versus central processing unit (CPU) cycles that are required by a single device per bit in wireless communication by either the cloud or edge computing. To achieve the latency requirements efficiently, the edge network should be designed by keeping in mind factors such as task complexity, processing power of the servers, and the network topology used. Regarding the bandwidth, edge computing reduces the data traffic by distributing the data among different edge servers for computational workload, and thus, lower amount of data is required to be shifted to the cloud.

The cooperation between cloud and edge computing provides high bandwidth as the bits are transmitted to the cloud server when the sum of the tasks exceeds the combined computational capacity of the edge servers. Another benefit of using edge computing in SG is the reduction in failure—if there is an electricity outage problem in a particular area of the grid, the edge computing services of the other areas will operate normally, without any problem. On the other hand, if the grid relies solely on cloud computing, and there is a power supply failure due to any natural disaster in the cloud infrastructure, then the whole network will fail [27]. As shown in the Figure 4, cloud computing shows the best performance when the signal to noise ratio (SNR) is low, but the edge computing performs the best even at high-SNR regime as the number of edge servers increases, outperforming the cloud-assisted counterpart.

The data from IoT devices that is not handled by the edge nodes are directly send to the cloud storage. The huge amount of data generated by these devices is stored in a low-cost storage at the cloud. In the second phase of data acquisition, the generated big data based on the volume, velocity, and variety is stored in a shared distributed fault

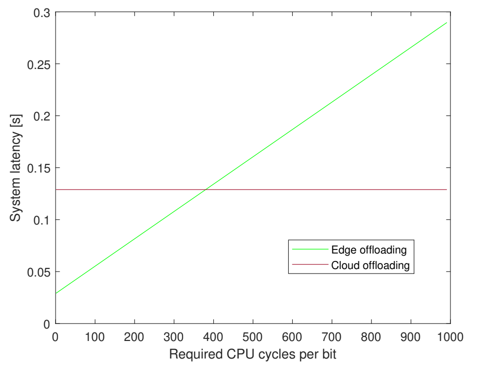



Fig. 3. System latency versus required CPU cycles per bit in wireless systems assisted by edge and cloud computing ($b = 1$ Mbit, $B^{\text{Edge}} = 10$ MHz $B^{\text{Cloud}} = 10$ MHz, $\gamma^{\text{Edge}} = \gamma^{\text{Cloud}} = 10$ dB, $f^{\text{Edge}} = 6$ GHz, $\tau^{\text{Cloud}} = 100$ ms) adapted from [4]

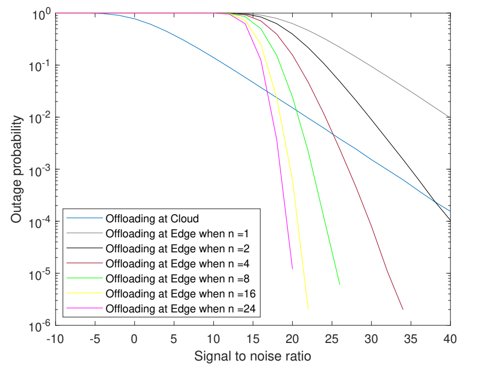



Fig. 4. Outage probability curves for wireless systems assisted by cloud and edge computing considering different numbers (n) of edge servers ($B^{\text{BH}} = 2$ MHz, $B^{\text{Edge}} = 50$ kHz, $R^{\text{target}} = 5$ Mbit/s) adapted from [4]

tolerant database. The collected data is then transferred into the master node(s) in the Hadoop cluster. As the data is collected from multiple heterogeneous devices, it may have different data formats and information, and therefore, data preprocessing will be required. In the data preprocessing, inaccurate and incomplete data are handled. Flume is used to perform the data acquisition process. The main function of Flume is to collect, aggregate, and transfer the large amount of data to Hadoop master node. The data received by Flume is stored in a single or multiple channel. The data is then sent to the external HDFS repository, where the data is written in a desired format using plug-in serializers. The serializers change and restructure the Flume data into the desired format. The data is preprocessed and a unified view of the data is achieved. The data is stored in the HDFS multiple clusters for processing. The HDFS clusters consists of DataNodes. The actual data and file system meta data are stored in those DataNodes. The data analysis is performed by the YARN on the data stored in HDFS; these two run on the same set of nodes that allows tasks to be processed on the nodes in which the SG data is present. Hive and Impala are the tools to perform SQL queries on data residing on HDFS. HIVE is used for data querying, to select, analyze and make calculations on the data of interest. The last phase is the data analytics; the tools used in Hadoop for data analytics is Scalable Advanced Massive Online Analysis (SAMOA), a distributed streaming ML framework consisting of programming abstraction for distributed streaming algorithms for data mining and ML tasks. Data visualization (graphs, reports, etc.) is done using Tableau, a common tool for interactive data visualization and sharing of information and dashboards.

## IV. Conclusions

The modern SG incorporates numerous heterogeneous devices. Due to this increase in the volume of structured, semi structured, and unstructured data, information retrieval from such a huge amount of data is a hard task. The collection, transmission, storage, processing, transformation, and analysis of large amount of data at a high rate are important for the efficient and effective function of modern SGs. The main aim of this research is to highlight the importance of IoT, edge computing, and big data for dealing with the high volume of SG data. In this paper, we have first presented the importance and requirements for big-data analytics in the SG. Subsequently, we have explained the applications of edge computing to the big data generated by the IoT devices in the SG. Edge computing is beneficial for SG in terms of latency, bandwidth, robustness to failure, and cost. In the future, we will apply big-data analytics to huge volumes of SG data and demonstrate the key requirements quantitatively.

## Acknowledgements

This paper is supported by Academy of Finland via: (a) ee-IoT project n.319009, (b) FIREMAN consortium CHIST-ERA/n.326270, and (c) EnergyNet Research Fellowship n.321265/n.328869.

## References

[1] A. Al-Ali *et al.*, "Role of internet of things in the smart grid technology," *Journal of Computer and Communications*, vol. 3, no. 05, p. 229, 2015.

[2] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[3] Y. Miao, G. Wu, M. Li, A. Ghoneim, M. Al-Rakhami, and M. S. Hossain, "Intelligent task prediction and computation offloading based on mobile-edge cloud computing," *Future Generation Computer Systems*, vol. 102, pp. 925–931, 2020.

[4] A. Narayanan, A. S. De Sena, D. Gutierrez-Rojas, D. C. Melgarejo, H. M. Hussain, M. Ullah, S. Bayhan, and P. H. Nardelli, "Key advances in pervasive edge computing for industrial internet of things in 5g and beyond," *IEEE Access*, vol. 8, pp. 206 734–206 754, 2020.

[5] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE internet of things journal*, vol. 3, no. 5, pp. 637–646, 2016.

[6] Stackpole Beth, "How industrial IoT is transforming the plant floor," 2016. [Online]. Available: https://searcherp.techtarget.com/feature/How-industrial-IoT-is-transforming-the-plant-floor

[7] Y. Ai, M. Peng, and K. Zhang, "Edge computing technologies for internet of things: a primer," *Digital Communications and Networks*, vol. 4, no. 2, pp. 77–86, 2018.

[8] F. Dong, J. Yong, and X. Fei, "Emerging intelligent big data analytics for cloud and edge computing," 2020.

[9] S. Choudhury, Q. Ye, M. Dong, and Q. Zhang, "IoT Big Data Analytics," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[10] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I. A. T. Hashem, A. Siddiqa, and I. Yaqoob, "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges," *IEEE Access*, vol. 5, pp. 5247–5261, 2017.

[11] C. Okoli and K. Schabram, "A guide to conducting a systematic literature review of information systems research," 2010.

[12] A. A. Munshi and Y. A. Mohamed, "Cloud-based visual analytics for smart grids big data," *2016 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2016*, 2016.

[13] S. S. Reka and T. Dragicevic, "Future effectual role of energy delivery: A comprehensive review of Internet of Things and smart grid," *Renewable and Sustainable Energy Reviews*, vol. 91, no. March, pp. 90–108, 2018. [Online]. Available: https://doi.org/10.1016/j.rser.2018.03.089

[14] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid — the new and improved power grid: A survey," *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 944–980, 2012.

[15] M. Ullah and K. Smolander, "Highlighting the Key Factors of an IoT Platform," *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 901–906, 2019.

[16] I. Lee and K. Lee, "The internet of things (iot): Applications, investments, and challenges for enterprises," *Business Horizons*, vol. 58, no. 4, pp. 431–440, 2015.

[17] M. Ullah, P. H. Nardelli, A. Wolff, and K. Smolander, "Twenty-one key factors to choose an iot platform: Theoretical framework and its applications," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10 111–10 119, 2020.

[18] A. A. Munshi and Y. A. Mohamed, "Cloud-based visual analytics for smart grids big data," in *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2016, pp. 1–5.

[19] S. Sagiroglu, R. Terzi, Y. Canbay, and I. Colak, "Big data issues in smart grid systems," *2016 IEEE International Conference on Renewable Energy Research and Applications, ICRERA 2016*, vol. 5, pp. 1007–1012, 2017.

[20] B. P. Bhattarai, S. Paudyal, Y. Luo, M. Mohanpurkar, K. Cheung, R. Tonkoski, R. Hovsapian, K. S. Myers, R. Zhang, P. Zhao, M. Manic, S. Zhang, and X. Zhang, "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions," *IET Smart Grid*, vol. 2, no. 2, pp. 141–154, 2019.

[21] G. Suciu, V. Suciu, A. Martian, R. Craciunescu, A. Vulpe, I. Marcu, S. Halunga, and O. Fratu, "Big Data, Internet of Things and Cloud Convergence – An Architecture for Secure E-Health Applications," *Journal of Medical Systems*, vol. 39, no. 11, 2015.

[22] E. Al Nuaimi, H. Al Neyadi, N. Mohamed, and J. Al-Jaroodi, "Applications of big data to smart cities," *Journal of Internet Services and Applications*, vol. 6, no. 1, pp. 1–15, 2015.

[23] J. Jin, J. Gubbi, T. Luo, and M. Palaniswami, "Network architecture and QoS issues in the internet of things for a smart city," *2012 International Symposium on Communications and Information Technologies, ISCIT 2012*, pp. 956–961, 2012.

[24] Z. Ding, X. Gao, J. Xu, and H. Wu, "IOT-StatisticDB: A general statistical database cluster mechanism for big data analysis in the internet of things," *Proceedings - 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCom 2013*, pp. 535–543, 2013.

[25] A. Shukla, S. Chaturvedi, and Y. Simmhan, "RIoTBench: An IoT benchmark for distributed stream processing systems," *Concurrency Computation*, vol. 29, no. 21, pp. 1–22, 2017.

[26] A. A. Munshi and Y. A. Mohamed, "Big data framework for analytics in smart grids," *Electric Power Systems Research*, vol. 151, pp. 369–380, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.epsr.2017.06.006

[27] J. Zhang, H.-W. Lee, and E. Modiano, "On the robustness of distributed computing networks," in *2019 15th International Conference on the Design of Reliable Communication Networks (DRCN)*. IEEE, 2019, pp. 122–129.

**ACTA UNIVERSITATIS LAPPEENRANTAENSIS**

**1079.** HUSSAIN, HAFIZ MAJID. Heuristic-based packetized energy management for residential electricity demand. 2023. Diss.

**1080.** HÄMÄLÄINEN, MINNA. Principals managing entrepreneurship education in schools. 2023. Diss.

**1081.** WANG, ZHAO. Photocatalytic degradation of pharmaceutical and person care products (PPCPs) by commercial and synthesized catalysts under UV irradiation. 2023. Diss.

**1082.** LOHRMANN, ALENA. The water footprint of the global power sector: Status quo, challenges, and opportunities for tackling the global water crisis. 2023. Diss.

**1083.** PONOMAREV, NIKOLAI. A salt and alkali synergy for synthesising active carbons from lignin: porosity development and techno-economic assessment. 2023. Diss.

**1084.** AFANASYEVA, SVETLANA. Wind farm layout optimization: project economics and risk. 2023. Diss.

**1085**. VOSTROV, KONSTANTIN. Reduction of non-circulating bearing currents by electrical machine design. 2023. Diss.

**1086.** CARILLO MELGAREJO, DICK. Improving the design of cellular networks beyond 5G for smart grids. 2023. Diss.

**1087.** MALACINA, IRYNA. Zooming in and out: The strategic management of innovation across multiple levels of complex supply networks. 2023. Diss.

**1088.** SORE, SARISEELIA. Impact of information system capabilities on business value creation: Aspects of IT-producing and IT-consuming companies. 2023. Diss.

**1089.** IMMONEN, EERO. Advances in optimal design of shapes and processes by computational fluid dynamics. 2023. Diss.

**1090.** LAVOYE, VIRGINIE. Augmented reality in consumer retail: a presence theory approach. 2023. Diss.

**1091.** HÄRKÖNEN, KALEVI. Smart buildings in the green energy transition. 2023. Diss.

**1092.** TSYTSYNA, EVGENIYA. Motives, uncertainties, and imbalances in the evolution of a sustainable business ecosystem. 2023. Diss.

**1093.** JÄÄSKELÄINEN, ATTE. Business model innovation opportunities when news has become a public good. 2023. Diss.

**1094.** ADEDIPE, TAIWO. Atmospheric boundary-layer characteristics and their significance in wind energy and urban air quality assessment. 2023. Diss.

**1095.** SOSUNOVA, INNA. Model and guidelines for designing Internet-of-Things-enabled smart waste management system in smart cities. 2023. Diss.

**1096.** VUORELA, JYRI. Operative business development through system model and changing business structures. 2023. Diss.

**1097.** TRIAPITCIN, ILIA. Knowledge injection method for real-time decision support. 2023. Diss.

**1098.** RÄISÄNEN, OTTO. Open data in distribution network asset management. 2023. Diss.

**1099.** MATELA, MIKA. Procurement improvement in the public agency. 2023. Diss.

**1100.** SHAH, DIPAL. Quantification of synchronization. 2023. Diss.

**1101.** GHAFOURI, MEHRAN. Thermomechanical finite element simulation of welding, and elevated-temperature mechanical behaviour of high and ultra-high strength steels. 2023. Diss.

**1102.** NEUVONEN, RIKU. Numerical ductile fracture assessment of weldments in direct-quenched, ultra-high-strength steel. 2023. Diss.

**1103.** HUPPONEN, MARI. Long-term evolution of greenhouse gas emissions from municipal solid waste management. 2023. Diss.

**1104.** WANG, QI. Dynamic analysis and parameter identification for robotic manipulators. 2023. Diss.

**1105.** KIMPIMÄKI, JAAN-PAULI. From observation to insight: Computational abduction and its applications in sustainable strategy research. 2023. Diss.

**1106.** YIN, RUOCHEN. Research on key technologies for lightweight maintenance operations of the remote handling system for a fusion reactor. 2023. Diss.

**1107.** MUNIR, QAISAR. Designing printing parameters for geopolymer concrete prepared from construction and demolition waste and industrial side streams. 2023. Diss.

**1108.** ROHANI RAFTAR, HAMIDREZA. Assessment of weld root fatigue strength of load-carrying fillet welded joints using notch stress approaches and finite element analysis. 2023. Diss.

**1109.** SADIQA, AYESHA. Sustainable energy transition for Pakistan: Assessing the role of energy, water supply, social and gender equity dimensions. 2023. Diss.

**1110.** GHOREISHI, MALAHAT. The role of artificial intelligence in circular value creation: A conceptual framework and evidence from case studies. 2023. Diss.

**1111.** SUURONEN, JARKKO. On numerical implementation of α-stable priors in Bayesian inversion. 2023. Diss.

**1112.** PIILI, HEIDI. A conceptual, holistic framework to overcome limitations and constraints of design in laser based powder bed fusion of metals: Case novel separation and purification units. 2023. Diss.

**1113.** LOPATINA, ANASTASIIA. From alternative sources of cellulose to ultrafiltration membranes. 2023. Diss.

**1114.** MARTTINEN, KATI. Toward more sustainable supply management: Practices, determinants, and enablers for ensuring sustainability across multiple supply chain tiers. 2023. Diss.

**1115.** HU, HAO. Development of hybrid microdevices composed of Ni-Mn-Ga and silicon layers: fabrication and metrology. 2023. Diss.

**1116.** MARAIA, RAMONA. Bayesian likelihood for intractable data. 2023. Diss.

ISBN 978-952-412-030-2
ISBN 978-952-412-031-9 (PDF)
ISSN 1456-4491 (Print)
ISSN 2814-5518 (Online)
Lappeenranta 2023