A DATA-INTEGRATED EDGE COMPUTING TECHNOLOGY ROADMAP FOR INDUSTRIAL INTERNET OF THINGS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

ALP BAYAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

APRIL 2023

Approval of the thesis:

A DATA-INTEGRATED EDGE COMPUTING TECHNOLOGY ROADMAP FOR INDUSTRIAL INTERNET OF THINGS

Submitted by ALP BAYAR in partial fulfillment of the requirements for the degree of **Master of Science in Information Systems Department, Middle East Technical University** by,

Prof. Dr. Banu Günel Kılıç
Dean, **Graduate School of Informatics** ____________________

Prof. Dr. Altan Koçyiğit
Head of Department, **Information Systems, METU** ____________________

Assoc. Prof. Dr. P. Erhan Eren
Supervisor, **Information Systems, METU** ____________________

Dr. Kerem Kayabay
Co-Supervisor, **HLRS, University of Stuttgart** ____________________

Examining Committee Members:

Prof. Dr. Altan Koçyiğit
Information Systems, METU ____________________

Assoc. Prof. Dr. P. Erhan Eren
Information Systems, METU ____________________

Assoc. Prof. Dr. Oumout Chouseinoglou
Information Systems and Technologies, Bilkent University ____________________

Date: *17.04.2023*

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Alp Bayar

Signature :

ABSTRACT

A DATA-INTEGRATED EDGE COMPUTING TECHNOLOGY ROADMAP FOR INDUSTRIAL INTERNET OF THINGS

Bayar, Alp
MSc., Department of Information Systems
Supervisor: Assoc. Prof. Dr. P. Erhan Eren
Co-Supervisor: Dr. Kerem Kayabay

April 2023, 83 pages

As the volume and velocity of data produced from production systems increase, transmitting all the data to a central cloud for processing becomes costly in time and network resources such as bandwidth and energy. In critical Industrial Internet of Things (IIoT) environments where a business decision requires an immediate course of action, edge computing as a new paradigm brings faster response time by processing data near where it is generated. As paradigms such as edge computing emerge, there is a lack of research on guiding digitally transforming organizations to utilize emerging technologies strategically. This study aims to investigate applications of edge computing and evaluate them regarding business objectives, market trends, technologies used, and challenges. To achieve those, first, this study investigates state-of-the-art edge computing architectures by their objective and application domain. Then, the study applies data-driven technology roadmapping from emerging technology management literature to edge computing applications in IIoT. It uses topic modeling, a natural language processing approach for identifying market and technology trends. It extends data-integrated roadmapping literature by using it as a technology and social change assessment tool, by integrating data as a layer of the data-driven technology roadmap.

Keywords: Edge Computing, Technology Management, Industrial Internet of Things, Technology Roadmapping, Topic Modeling

ÖZ

ENDÜSTRİYEL NESNELERİN İNTERNETİ İÇİN VERİ GÜDÜMLÜ UÇ BİLİŞİM TEKNOLOJİ YOL HARİTASI

Bayar, Alp

Yüksek Lisans, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Doç. Dr. P. Erhan Eren

Tez Eş Danışmanı: Dr. Kerem Kayabay

Nisan 2023, 83 sayfa

Üretim sistemlerinden üretilen verilerin hacmi ve hızı arttıkça, tüm verinin işlenmek üzere merkezi bir bulut sunucusuna iletilmesi zaman, bant genişliği ve enerji gibi kaynaklar açısından maliyetleri artırmaktadır. Veri odaklı karar vermek adına, bir organizasyonda anlık olarak karar almak gereken kritik endüstriyel nesnelerin interneti ortamlarında, yeni bir bilgi işlem paradigması olarak uç bilişim, verileri üretildiği yere yakın bir yerde işleyerek daha hızlı yanıt süresi sağlar. Uç bilişim gibi modeller ortaya çıktıkça, dijital dönüşüm sürecinde olan kuruluşlara gelişmekte olan teknolojileri stratejik olarak kullanabilme konusunda rehberlik edecek araştırma ve bilgi eksiklikleri tespit edilmiştir. Buna istinaden bu çalışma, uç bilişim uygulamalarını araştırmayı ve bunları işletme gereksinimleri, pazar eğilimleri, kullanılan teknolojiler ve karşılaşılan zorluklar açısından değerlendirmeyi amaçlamaktadır. Bu doğrultuda ilk olarak; güncel uç bilgi işlem mimarisinin endüstriyel nesnelerin interneti uygulamaları araştırılmış, çalışmalar öncelikli amaçları ve uygulandıkları alana göre incelenmiştir. Daha sonra teknoloji yönetimi literatürü incelenmiş ve veri odaklı teknoloji yol haritası çerçevesi endüstriyel nesnelerin interneti ortamlarında uç bilgi işlem mimari ve uygulamaları konusuna uygulanmıştır. Pazar ve teknoloji eğilimlerini belirlemek için bir doğal dil işleme uygulaması olan konu modelleme algoritmaları kullanılmıştır. Teknoloji yol haritasının bir katmanı olarak veri, yol haritasına dâhil edilerek, veri odaklı yol haritası, bir teknoloji ve sosyal değişim değerlendirme aracı olarak genişletilmiş ve veri güdümlü yol haritası literatürüne eklenmiştir.

Anahtar Sözcükler: Uç Bilişim, Teknoloji Yönetimi, Endüstriyel Nesnelerin İnterneti, Teknoloji Yol Haritalama, Konu Modelleme

To My Family

ACKNOWLEDGMENTS

First of all, I would like to express my thanks to my advisor Assoc. Prof. Dr. P. Erhan Eren for his guidance, support, and all the opportunities he provided. His supervision allowed me to work in an environment along with talented researchers which enabled me to enhance my skills and knowledge. I would also like to thank my co-advisor Dr. Kerem Kayabay, who was always supportive, self-devoted, and patient throughout my Master's journey. He provided great interest, both academic and social guidance for this thesis and was a great mentor affecting my way of thinking as a researcher and work discipline in academia. Besides my supervisors, I extend my sincere thanks to Umut Şener, who also helped me a lot publish my first conference proceeding as a co-author, which is also part of this thesis, and also guided me to attend my first international conference in Europe.

I would also like to thank all of my colleagues and friends from Informatics Institute, Burcu Koç Göltaş, Kerem Nazlıel, Arif Ozan Kızıldağ, Mert Mecit, Yılmaz Taylan Göltaş, Volga Sezen, Ezgi Çavuş, Şeyma Şimşek, Cemre Yazıcı and Toyan Ünal for their companionship and support. Also, thanks to Ege Kılınçkaya, Anıl Erkul, Çağlar Sevinç, and Onur Aksoy for their friendship and collaboration in course projects. I also thank Dr. Mert Onuralp Gökalp, Bahadır Yandık, and Assoc. Prof. Dr. Ebru Gökalp for their pieces of advice. Special thanks to the administrative staff, Sibel Gülnar, Nilüfer Şeker, Hakan Güler, and Ayşe Nur Özdere Yüksel, for always being friendly, helpful, and cheerful.

Also, thanks to Prof. Dr. Altan Koçyiğit and Prof. Dr. Tuğba Taşkaya Temizel for the skills I have acquired from their lectures that made this thesis possible.

Finally, I would like to offer my gratitude to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting my expenses for attending the GECON 22 Conference in Izola, Slovenia.

I wish peace and prosperity to our country and a bright future for all my family and friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AGV	Automated Guided Vehicle
API	Application Programming Interface
AR	Augmented Reality
BERT	Bidirectional Encoder Representations from Transformers
CoT	Convergence of Cloud Computing and IoT
CPU	Central Processing Unit
c-TF-IDF	Class-Based Term Frequency-Inverse Document Frequency
CRISP-DM	Cross-Industry Standard Process for Data Mining
DL	Deep Learning
DNS	Domain Name System
ERP	Enterprise Resource Planning
GPT	Generative Pre-trained Transformer
HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
HCP	Hyperscale Cloud Provider
IT	Information Technology
IoT	Internet of Things
IIoT	Industrial Internet of Things
IoV	Internet of Vehicles
IP	Internet Protocol
KPI	Key Performance Indicator
LDA	Latent Dirichlet Allocation
LORA	Low Power Wide Area Network
LSI	Latent Semantic Indexing
MRI	Magnetic Resonance Imaging
ML	Machine Learning
MQTT	MQ Telemetry Transport
NFV	Network Function Virtualization
NLP	Natural Language Processing
OT	Operational Technology
PCA	Principle Component Analysis
PLC	Programmable Logic Controller

pLSI	Probabilistic Latent Semantic Indexing
QFD	Quality Function Deployment
RAN	Radio Access Network
RFID	Radio-Frequency Identification
SAO	Subject Action Object
SCADA	Supervisory Control and Data Acquisition
SDN	Software Defined Networking
SLR	Systematic Literature Review
SME	Small and Medium-Sized Enterprise
SWOT	Strengths, Weaknesses, Opportunities, Threats
TF-IDF	Term Frequency-Inverse Document Frequency
TRM	Technology Roadmap
TRT	Technology-Relationship-Technology
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solutions
TSN	Time Sensitive Network
t-SNE	t-Distributed Stochastic Neighbor Embedding
UAV	Unmanned Aerial Vehicle
UMAP	Uniform Manifold Approximation and Projection
VR	Virtual Reality
VNF	Virtual Network Function

CHAPTER 1

INTRODUCTION

1.1 Research Background

In digital transformation, improvements in digital technologies enable people to build new capabilities and improve the performance of applications in business use cases, and these result in social changes driven by changes in the competitors and customers (Pousttchi et al., 2019; Vial, 2019). Although different definitions and classifications exist; cloud computing, internet of things, big data, and data analytics are a few examples of digital technologies (Pousttchi et al., 2019), and paradigm changes evoked by them enable the Industry 4.0 revolution (Zaki, 2019). In the context of Industry 4.0, machines in a manufacturing environment can interact with their environment, and are context-aware, smart, and even self-learning (Gokalp et al., 2016). Enabler technologies of this paradigm either can be new cyber-physical technologies, or some of them refer to the convergence of multiple technologies and some are only paradigm shifts in using available technologies in a new business model. These data-driven business models enabled by digital technologies are shaped around data (Fruhwirth et al., 2020), and are concerned with collecting, transmitting, processing, and using it as efficiently and as effectively as possible to digitally transform and achieve data-driven organizations.

The Internet of Things (IoT) brought many new opportunities to industrial systems. The term coined by Kevin Ashton in 1999, is the idea of enabling objects to manage their tasks without human intervention by communicating with each other (Ashton, 2009). These objects generally see the world through the data gathered from their sensors and actuators and can manage their tasks without human intervention. Industrial IoT (IIoT) refers to IoT applications in the industrial domains. According to several market research companies (AllTheResearch, 2021; MarketsandMarkets, 2021), the IIoT market is around $77 billion and is expected to reach around $110 billion by 2027. Gartner predicts that around 75% of business data will be generated outside a traditional central cloud architecture by 2025. This number was around 10% in 2021 (Gartner, 2018).

Data from production systems are traditionally sent and processed in cloud-based central computing architectures (Oracle, 2023). As the volume and velocity of data increase, transmitting all the data to a central cloud for processing becomes costly in time and network resources such as bandwidth and energy consumption. In critical IIoT environments where a business decision requires immediate computation and action, edge computing as a new paradigm brings faster response times (Yi et al., 2015). The term “edge” refers to the nodes in the opposite direction of the cloud data center (Accenture, 2023). Rather than transmitting and processing a high volume of data from IIoT devices to a central cloud, edge computing reduces network congestion by moving computing and storage near where data is generated (Qiu et al., 2020). This reduced latency enables real-time processing of data and enables many opportunities to improve an organization's business processes. In an edge architecture, sensors, actuators, and other production machine logs can be sent to edge nodes, which are computing devices with smaller capabilities compared to a cloud server. These edge nodes can analyze, process or store the whole or a fragment of data. There are many edge architectures proposed, varying in technological properties depending on the business requirements of the use case they are applied (Yu et al., 2018).

Investigating available real-life applications and understanding edge architectures is significant for organizations since practitioners can understand the technology trends and architectures for a particular IIoT edge computing use case. In the following section, problems and lack of knowledge identified from the socio-technical perspective on edge computing applications in the literature are questioned.

1.2 Problem Statement

Analyzing technological development patterns and convergences, and predicting future technologies and their relations to market trends are important questions of research in many fields, helping organizations to achieve competitive advantage (Zaki, 2019). Practitioners and researchers are increasing awareness of the potential of edge computing, however, the type of architectures to use in their case to achieve which business objective is a question, also there are many challenges in utilizing those demo applications in production and slowing digitally transforming the business processes (Atieh, 2021). Although edge computing technology makes a good fit to use in IIoT systems, the data underachiever nature of the manufacturing industry (Kayabay et al., 2022) and lack of strategic technology management in industrial corporations makes it hard to adopt a cloud-edge computing best practice architecture (Bayar et al., 2023). While paradigms such as edge computing emerge, there is a research gap in edge computing literature, on guiding digitally transforming organizations to utilizing emerging technologies strategically. Current studies review existing edge architectures, possible application domains, and objectives; however, they lack a business strategy perspective.

Technology roadmapping is a tool for technology management and strategic technology planning. It can be used for exploring and understanding relationships between markets, technological resources, and strategic objectives, also product, service, strategy planning, and more (Phaal et al., 2004b). It is a flexible and customizable tool, which can have different goals such as planning, forecasting, assessment, and scopes such as industrial, national, and organizational. At the sectoral level, it helps to identify and forecast future trends using exploratory methods while at an organizational level, it increases communication between stakeholders for planning markets, products services and technologies over time (Lee and Park, 2005).

Indeed, technology management tools such as roadmapping can help practitioners understand technology trends and architectures (Lee and Park, 2005) but no study explores any technology management framework for edge computing in IIoT. Nonetheless, while developing a strategy for a computing paradigm, an important asset to consider is the data, subject to be transferred, processed, analyzed, or stored for the business processes (Han and Geum, 2020). New tools and frameworks for management and strategy should be shaped considering the value of data, as the main resource of digital transformation. (Han and Geum, 2020) integrated data layer to a technology roadmap for smart service planning. (Kayabay et al., 2022) proposed a data science roadmapping framework for organizations to structurally plan a digital transformation process in order to become data-driven. The study uses roadmapping as an organizational planning tool, conducting workshops for data-related, technological, and organizational resources, and discusses it can be extended as a social assessment and evaluation tool to complement the data-integrated roadmaps for future research. Based on this, it is evident that there is a lack of using quantitative approaches, especially with the rapid rise of transformer-based NLP algorithms, in large textual databases for developing data-integrated strategic planning and technological assessment for emerging technologies. This thesis uses roadmap as a social assessment and evaluation tool and discusses it can be used with other technology forecasting tools and can be integrated to complement the data-integrated roadmaps for future research as discussed in (Kayabay et al., 2022).

1.3 Research Aim and Objectives

This study aims to investigate applications of edge computing and compile academic studies regarding business objectives, market trends, technologies used, and challenges. The findings of the study will provide insights into edge computing adoption, research, and development. It will increase the understanding of organizations and practitioners. Research objectives are stated below:

I. Investigating applications of edge architectures in industrial environments to understand the motivations behind using these technologies.

II. Identifying and mapping the relationships between technology and market trends of edge computing.

III. Developing a data-integrated edge computing technology roadmap using data-driven approaches.

1.4 Significance of the study

This study integrates research areas of edge computing and data-driven roadmapping for emerging technologies.

I. It investigates edge computing applications in IIoT and evaluates them by their business objectives and application domains.

II. It investigates data-driven technology roadmapping literature to bridge the sociotechnical knowledge gaps in the edge computing domain.

III. Methodologically, it contributes to exploratory technology assessment and forecasting literature by combining SLR and topic modeling into technology roadmapping.

IV. It extends data-integrated roadmapping literature as a technology and social change assessment tool and applies it to the edge computing domain.

V. It integrates dynamic topic modeling into technology roadmapping stated as in future work by (Feng et al., 2022; Kim and Geum, 2021). It evaluates and compares BERTopic and LDA algorithms for creating topic models.

1.5 Structure of the Thesis

The subsequent parts of this thesis are organized as follows: After the introduction in Chapter 1, Chapter 2 presents a comprehensive review of literature in related research areas, and evaluates the state of knowledge. As stated in Chapter 2.10, knowledge gaps identified in the literature review revealed the need for this study. Edge computing literature and socio-technical studies on other emerging technologies are investigated to bridge two research areas. Chapter 3 presents the research aim, research questions, research objectives and chosen methods to be used for achieving the objectives of the study. As one of the chosen methods, the systematic literature review is presented in Chapter 4 which examines edge computing applications in IIoT. Chapter 5 employs quantitative exploratory methods for identifying market and technology trends and developing an edge computing technology roadmap for IIoT. Finally, Chapter 6 concludes the study by summarizing and discussing findings, results, limitations, and prescribes fields of future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Edge Computing and Related Terms

In this chapter, a short history of edge computing and related fog computing, multi-access edge computing, mobile edge computing, and cloudlet terms are explained. Practically, these terms can be used interchangeably in combinations. However, conceptual definitions and how terms are used in literature de facto are explained below.

2.1.1 Edge Computing

As introduced in Chapter 1.1, edge computing is a distributed computing paradigm referring to processing data closer to sources, increasing processing speeds, enabling to handle of high volumes of data without causing bandwidth congestion, and yielding a lower response time to end-users (IBM, 2020). The origins of edge computing can be traced back to the content delivery network proposed by Akamai in the early ’90s (Kuever, 2019), to deliver multimedia content to users using caching in edge servers closer to end-users (Dilley et al., 2002). Another important milestone for edge computing was (Noble et al., 1997), conducting speed recognition using offloading from mobile devices with limited computing capability to central servers. These studies were followed by academic research on improving battery life and improving computing capabilities of mobile devices (Satyanarayanan, 2017). Technical research focuses on improving technical properties of edge architectures such as security issues, the energy consumption of hardware, computational limitations and efficient offloading, and limited storage capacity (Bayar et al., 2023).

Motivations of using edge computing in IIoT environments are explained in Chapter 2.2, followed by challenges of implementing edge architectures in Chapter 2.3, open-source tools available in the edge landscape in Chapter 2.4, and market roles and responsibilities in Chapter 2.5.

2.1.2 Fog Computing

Fog computing is a similar concept introduced by Cisco, focusing more on the infrastructure between edge devices and central cloud servers (Qiu et al., 2020). After reviewing studies in the literature, Fog and Edge terms are used interchangeably in this study because what they are trying to achieve is the same from the business objective perspective. The fog computation term is used in the literature that focuses on the architecture and technical infrastructure level and is considered a subset of edge computing (Qiu et al., 2020).

2.1.3 Cloudlet

Cloudlet terms were introduced by (Satyanarayanan et al., 2009), which propose the core feature of edge computing, reducing latency using a data center in a box connected to the internet, rather than using a central cloud. Cloudlets reside in a middle layer between the mobile edge device and the central cloud, and function as a small data center providing cloud capabilities closer to mobile devices. In 2017, NIST defined a cloudlet as a virtual fog node in a fog architecture similar to a virtualized switch, or a virtual machine (Iorga et al., 2018).

2.1.4 Mobile Edge Computing

This term refers to networks with mobile devices at the edge, namely smartphones, and tablets. Initially, ETSI used the MEC acronym for mobile edge computing, but as the research progressed, they realized edge devices, mostly IIoT devices in manufacturing do not have to be mobile devices. So in 2017 ETSI officially changed the term in the research group's name from Mobile Edge Computing to Multi-Access Edge Computing (Dahmen-Lhuissier, n.d.). The core of Mobile edge networks is built on virtualization. European 5G Public Private Partnership defines MEC as A key enabler for 5G networks together with NFV and SDN technologies (Hu et al., 2015).

2.1.5 Multi-Access Edge Computing (MEC)

MEC is an extension of mobile computing in edge computing. ETSI defines the MEC as a technology providing IT and cloud computing services within Radio Access Network (RAN) in 4G and 5G (Fabio Giust and Gianluca Verin, 2018). MEC extends edge computing, reducing computing and storage energy consumption and bringing mobile devices with lower calculation capabilities when compared to cloud computing infrastructure. ETSI defined five key features of MEC was on-premises, proximity, reduced latency, location awareness, and network context information (Carvalho et al., 2021). Although there are technical differences between MEC, cloudlet, and fog computing, they are all considered under the edge computing umbrella in this study.

2.2 Edge Computing in IIoT: Motivation and Considerations

Edge computing has been a complementary and supporting paradigm to cloud computing rather than competitive. For understanding what edge computing brings, in an experiment carried out by (Yi et al., 2015) on a face recognition task, it has been shown that moving applications from the cloud to the edge reduces 900 msec to 160 in response time.

With better response times, edge nodes can help monitor and control processes (L. Li et al., 2018) or machine status (Bose et al., 2019), make forecasts under uncertainty (Taïk and Cherkaoui, 2020), spot bottlenecks, inefficiencies, and deep-dive failures into their root causes. There are studies in the following sections that will provide information about how these decisions are made in literature while communicating machines with edge architectures keeping optimal bandwidth, energy consumption, costs, and processing capabilities.

In terms of data security and privacy, in an edge architecture, data is less transmitted in a network, therefore, is less exposed to other network components, reducing the single point of failure in central architectures and reducing risks of data leakages (Qiu et al., 2020). However, vulnerabilities are higher as stated in the section below, differences in edge nodes can cause and might bring different security requirements than in cloud architectures (Carvalho et al., 2021).

2.3 Edge Computing in IIoT: Challenges

Unlike central cloud systems, edge computing systems are decentralized, consisting of different technologies on the edge nodes, so resource management needs to be worked on more. In their review, (Khan et al., 2019) state that edge computing platforms are heterogeneous considering data communication, protocols, APIs, policies, platforms, and energy consumption.

In a similar study to this, (Qiu et al., 2020) present current challenges in literature such as security, routing, task scheduling, energy efficiency, data storage, analytics, and standardization.

Several storage servers may be deployed with different operating systems to store and manage data, which creates naming problems. Traditional naming systems like DNS are insufficient for dynamic edge computing networks, and IP-based naming becomes too costly for edge nodes with multiple sources and tasks (Yu et al., 2018).

As computing shifts from the center to the edge, scalable and elastic sharing of resource pools becomes hard to achieve and is limited. Edge devices are harder to manage the lifecycle and orchestrate dependencies between them. Edge computing needs more effort in infrastructure management mainly due to the heterogeneity of devices, different communication protocols, technologies, and several other constraints like bandwidth, CPU, memory, and battery. There are standardization efforts for orchestration, such as ETSI MEC and ETSI NFV Management and Orchestration (Industrial Internet Consortium et al., 2016).

2.4 Edge Computing Landscape and Open-Source Tools

Working together with academia and companies with areas of expertise is vital for achieving the best practices of IIoT. Being part of initiatives, contributing to consortiums, and helping collaborations mentioned in this chapter are essential. The OpenFog Consortium was founded in 2015 by companies like Cisco, Intel, Dell, and Microsoft. The consortium merged with the Industrial Internet Consortium in 2019, which also had similar efforts, for example, a reference architecture called OpenFog, which many studies used as a reference.

EdgeX Foundry is an open-source software framework, one of the eight projects hosted by the LF Edge organization of the Linux Foundation. LF Edge projects aim to create a unified edge community for increasing collaboration and cooperation. EdgeX provides a vendor-neutral, open-source IIoT edge computing common framework (The Linux Foundation, 2021). (Liu et al., 2019) reviewed and explained the differences between open-source edge computing tools considering performance, applicability, and energy-efficient deployment strategies. Liu’s work can guide the selection of appropriate open-source tools for a company to build its edge infrastructure. It is stated (Liu et al., 2019) that EdgeX is designed for IoT environments with various sensors, making it practical in automated factories, machinery systems, and industrial use cases. Another LF Edge project is Akraino, a set of integrated infrastructure that can serve a broader range of use cases. According to (Liu et al., 2019) it is created for network operators who focus on the edge infrastructure. One of the projects in an earlier stage, KubeEdge is worth mentioning for extending containerization orchestration capabilities into edge architectures. StarlingX by OpenStack, an edge infrastructure software platform (StarlingX, 2021), Eclipse Kura, ioFog, and fog05 platforms enable people to build IIoT gateways.

Some of the frequently encountered open-source tools used in the architectures in the literature are explained below: Modbus, OPC UA, MQTT, and ZigBee are examples of open protocols frequently mentioned in open edge architectures literature. EdgeX can communicate with devices using these protocols, making it a good fit for IIoT. MySQL and MongoDB are popular open choices for storage. Grafana and Ganglia are monitoring and visualization systems. Docker is a popular containerization tool. Containerization is a natural fit for edge computing due to easy deployment and light storage and computing requirements. Kubernetes is an orchestration tool for managing containerized applications in a scalable manner.

2.5 Edge Computing Market Players and Roles

According to Ericsson's report (Carlos Brava and Henrik Backström, 2020), companies that operate in the domain are traditionally Cloud Service Providers, OT vendors, System Integrators, and Communication Service Providers. Companies like AWS, Microsoft, and Google are mentioned as Hyperscale Cloud Providers (HCP) in the paper as other key players in the edge ecosystem. With their bountiful resources, they can serve many companies from a variety of sectors and are able to drive the ecosystem. OT vendors are companies that utilize IoT as their primary business role. Companies such as Siemens, GE, and BMW use OT for their activities in smart manufacturing. These companies are more likely to coordinate with Cloud providers to realize applications and develop best practices. System Integrators are providers of edge computing tailor-made specific solutions that address varying market needs. They can also serve with tailor-made specific solutions.

An edge computing market player can serve as a combination of full edge provider, partner edge provider, aggregator edge provider, and limited edge provider (Carlos Brava and Henrik Backström, 2020). Full edge providers provide infrastructure and platforms. Partner edge providers help with connectivity and reconfiguration of present infrastructure. Aggregator edge providers provide infrastructure software and platform as a service, and limited edge providers cover enterprises for most of their edge computing requirements, deciding edge computing setup. It is essential to know these roles and strategies in the market while selecting a partner to achieve company goals and forming an edge computing adoption strategy.

2.6 Edge Computing as Enabler of IT/OT Convergence

Traditionally, Information Technology (IT) is used for a more centralized office environment, computers, storage, networking devices, and infrastructure to process data. Operational Technology (OT), on the other hand, is associated with the edge devices in manufacturing like programmable logic controller (PLC), supervisory control, and data acquisition (SCADA) that were traditionally not connected to the internet, used to manage industrial processes or equipment (Denzler et al., 2020). IT/OT convergence has been discussed for standardizing OT, providing IIoT data to business strategic decision-making, and enabling new strategies and business models to utilize data. In IIoT, OT devices are generally connected as a distributed network architecture, making edge computing more convenient for the following: devices have connectivity bottlenecks, applications need large data transmission, preprocessing is required before transferring data, or there are security issues for transferring all the data to a center for processing (Pop et al., 2021).

By reengineering themselves in these ways, organizations will have a more convenient collection and analysis of the data obtained from a network of edge devices. IT/OT convergence reflects the true potential of building an edge computing architecture, faster decision-making in the industry, and increased business value through data-driven operations.

2.7 Edge Computing in IIoT Literature

Edge computing applications in IIoT domain are to be investigated, focusing on studies with similar methodologies such as surveys and literature reviews. These secondary studies include information on the research focus such as creating a unified edge architecture, supporting technologies, technical opportunities, and challenges.

(Banijamali et al., 2020) conducted a systematic literature review to identify critical architectures used in the convergence of cloud computing and IoT, referred to as CoT. In the search strategy, they included edge or fog or cloud keywords in the search strings, which is why this study is related to our research. Also examining the relationship between the architectural style and quality attributes, they found performance efficiency, portability, and security have the most research attention.

As their results suggest, the number of manufacturing applications was behind after smart cities, transportation and mobility systems, healthcare, and smart homes holding only 2% of the architectures in their research were in manufacturing. Investigating the popularity of "architectural design patterns" authors find "edge connectivity" pattern was the first followed by stream processing, virtual device representation telemetry ingestion. In edge connectivity, IoT gateways can connect any kind of device to the cloud without any adaptation problem (Banijamali et al., 2020). Although this work is comprehensive, industrial practitioners need more specific detailed guidance, and as stated in the future work section, standards or frameworks are needed for developers and researchers to select from existing technologies or develop their tailored solutions.

(Khan et al., 2019) conducted a comprehensive survey on edge computing. The study discusses the edge computing paradigm, the latest problems, and use cases in businesses where low response time is essential. Classified edge computing as fog, cloudlet and mobile edge architectures and compared papers in these classes based on the objectives identified for each class. Following, identified key requirements for edge computing, such as smart billing mechanism, real-time applications support, a joint business model for management and deployment, resource management, resource management, scalability, security and reliability capabilities. This study also evaluates edge solutions by their objectives.

In a similar study, (Qiu et al., 2020) review surveys of edge computing, their focuses, findings, and differences, and discuss that the edge computing architectures in the literature do not explain each layer in detail. Then present current challenges in literature such as security, routing, task scheduling, energy efficiency, data storage, analytics and standardization. They suggest their proposed reference architecture differs in terms of considering the characteristics of IIoT in layers in more detail. Possible application areas of the architecture are also discussed such as smart vehicles, grids and manufacturing. Then, opportunities and challenges in the widespread adoption of edge computing are given.

(Hamm et al., 2019) reviewed edge computing initiatives worldwide, and characterized them using layers that focus on edge computing. Then proposed a roadmap for edge computing development in terms of social, ecological, and economic sustainability, relating them to the main concerns of edge computing introduced in Reference Architecture Model for Edge Computing (RAMEC). These concerns are security, real-time capabilities, smart capabilities, and management capabilities.

(Carvalho et al., 2021) state there is no consensus on the terms fog, edge, mobile edge, cloudlet, MEC and give examples of different views in the literature. They map studies with use cases and these paradigms with a table. Also, they present a chronologic examination of edge computing research topics, which is useful for researchers and practitioners to follow state-of-the-art applications.

As explained below under Chapter 4, related studies focus on the technological aspects of edge architectures. They explain application domains, when and why to use technologies, and further work on the open research topics, however, most of the applications are experimental. It is found that there is a lack of a strategic deployment decision framework for the implementation of such technologies. All applications are trial-based, without sufficient information about organizational research and feasibility. After identifying the need for more strategic and sociotechnical knowledge on these issues, technology management literature is investigated for studies applied to similar different emerging technologies.

2.8 Technology Management Activities

Technology management activities along with supporting activities are explained to avoid confusion of terms. Management of Technology provides a hidden competitive advantage by bridging the knowledge and practical gap between science, engineering and management (National Research Council, 1987).

The core TM processes for managing and building technological capabilities: (Cetindamar and Phaal, 2017; Gregory, 1995; Rush et al., 2007)

1. Identification of technologies that are currently not employed by the organization but might be necessary.

2. Selection of technologies in line with business strategy.

3. Acquisition of selected technologies, make-buy decisions.

4. The exploitation of technology, utilizing new capabilities into business value.

5. Protection of knowledge and know-how (e.g. patenting).

6. Learning from all the knowledge produced in the process and making the TM processes sustainable.

There are also activities supporting TM activities. Knowledge management, project management, and innovation management:

Knowledge management aims to add and create value by leveraging know-how, experience, and judgment within and outside an organization. Comprises a range of practices used by organizations to identify, create represent and distribute knowledge for reuse, awareness, and learning. Knowledge includes awareness, cognition or recognition (know-what), capacity to act (know-how) as well as understanding (know-why) that is within the mind (Desouza, 2005).

Innovation management is the successful implementation of novel ideas in the form of a product, service, or business process. Innovation does not have to be limited to new technology, it can be a new business model, such as the cloud computing paradigm itself. Technology Management includes all decision-making to develop or use a technology within an organization (Cetindamar and Phaal, 2017).

2.9 Technology Management Tools

In this study, a method, tool, or framework for answering the problems stated in Chapter 1.2 regarding the edge computing domain is investigated from the technology management domain. Here, related alternatives are explained briefly.

2.9.1 Technology Roadmapping

Although there are older appearances of technology roadmapping by organizations such as NASA, Boeing, GE, and Lockheed (Kerr and Phaal, 2020); the first published journal paper in manufacturing by Motorola in the late 80s integrated product and technology strategic planning (Willyard and McClees, 1987). Since then, roadmapping is widely used for supporting business strategy, innovation and policy. Technology roadmapping is a technology management tool and a process, while the output of the process is the technology roadmap. Roadmap itself is a visual representation having varying formats. One of the most generic roadmaps used is a layered chart with market, product, and technology layers with a time axis (Phaal et al., 2004b). A popular fast-start method also called the "T-Plan", developed by the Cambridge practical school, provides a starting point for organizations by conducting several workshops focusing on different aspects. The roadmapping process is based on the organizational and case-dependent context therefore can be customized along with the roadmap structure (Phaal et al., 2004a).

As explained in the problem statement in Chapter 1.2, technology roadmaps can be used for product planning, service planning, strategic planning, foresight, knowledge planning, technology forecasting, and technology assessment. Depending on the different purposes, the typology of the roadmaps can also be modified. The type of technology roadmapping process for mapping technologies to business operations and capabilities as service/capability planning (Phaal et al., 2001). This typology of technology roadmap is ostensibly convenient for the research questions of this study. In terms of time, two common types of technology roadmapping are retrospective and prospective analysis, also called backward and forward (Kostoff and Schaller, 2001). Retrospective analysis covers time frames from past to present using existing data and has a higher level of certainty, objectivity, and reliability then prospective analysis (Kostoff and Schaller, 2001). Extending the roadmapping time horizon to the past enables us to understand industrial emergence and map scientific and technological developments (Phaal et al., 2011). Analyzing how academic knowledge changes over time retrospectively and how industries change enables us to analyze innovations of socio-technical transitions (Park et al., 2020).

Roadmapping studies can also be grouped by the methods used for creating the roadmaps. (Pora et al., 2022) generalizes these methods as expert-based, computer-based, and hybrid methods. This categorization is similar to comprehending qualitative and quantitative methods since most of the quantitative approaches are executed using computers. The aforementioned fast-start T-Plan method is an example of an expert-based method, which utilizes knowledge from concluding workshops between stakeholders. There are also schools of thoughts on roadmapping (Park et al., 2020) where different methodologies used in roadmapping studies are grouped by different universities around the world are classified. For example, Cambridge's practical school, mentioned above, generally prefers expert-based methods, while Seoul school prefers more computer-based methods, which are investigated in detail in the following chapter.

2.9.2 Data-Driven Roadmapping

There are several methods to collect data for generating roadmaps when workshop-based roadmapping is not possible or not preferred. This is generally the case in studies of practitioners classified as the Seoul school (Park et al., 2020). Large textual databases including technology, science, engineering, and product content can be analyzed using approaches such as bibliometric methods, co-occurrence or co-citation analysis, and patent analysis in a both retrospective and prospective way (Kostoff and Schaller, 2001). In emerging technologies, where there is no historical data available (Daim et al., 2006) state bibliometric and patent analysis can be used as data collection methodologies for technology forecasting tools such as scenario planning or growth curves. Also, Quality Function Deployment, Bayesian Networks, several text mining approaches, and patent analysis are other quantitative methods that are used in technology roadmapping studies. Below, examples from the literature are explained with a focus on the methodologies they used for developing a roadmap. All studies are evaluated and grouped by data sources and methodologies they used for identifying roadmap elements and defining relationships between those elements, information on the typology of the roadmap and presented in Table 1.

(Zhang et al., 2013) conducted co-occurrence analysis and frequency analysis on publications and patent records as bibliometric methods and constructed a technology roadmap accordingly, using feedback from experts. As a case study, they worked on electric vehicles which were in the very early stages as an emerging technology of that time. In another similar study, (Zhang et al., 2014) worked on Dye-Sensitized Solar Cells. The study explains the foundations of Triz, the theory of inventive problem solving proposed in the 1940s, used for extracting similar ideas from massive numbers of patents, and how Subject-action-object analysis is integrated Triz to make it semantic. When compared to PCA or LDA, authors stated Triz needs more qualitative effort (such as expert opinion) than LDA, an advantage of Triz is understanding problem and solution patterns, but as a disadvantage, it is dependent on qualitative assessment. Notice NLP applications in these years did not have their performance of today, which can be identified when looking at the methods used as relatively new studies.

(Jin et al., 2015) propose a technology-driven roadmap instead of a market-driven one for solar LED lighting technologies using text mining and similarity scores. Markets and products are evaluated based on technologies available and developing. Patents and product user manuals are used as data sources. They discussed in future work, TF-IDF scores and SAO analysis can be used for overcoming imitations of their methodology. Two authors of that study (Jeong and Yoon, 2015) created a patent roadmap using text mining and patent analysis, grouped and classified patents using qualitative patent analysis methods such as TEMPEST, and identified patterns for patent planning for AMOLED display technology.

Regarding the relationship between layers of the technology roadmap, (Geum et al., 2015) used association rule mining for measuring the dependencies of keywords between layers. In a case study they collected data about Apple products and services from the internet for creating data sources for product layer and service layer. They use text mining from documents to identify keywords of these layers. Using association rule mining, they created a keyword portfolio roadmap and keyword relational map. To investigate the relationship between elements (Son and Lee, 2019) proposed using network analysis with fuzzy-set theory and presented a case study in 3D printing technology. Elements of the roadmap are chosen after the expert workshop.

Topic modeling with LDA was performed on patent data for technology assessment and roadmapping of another emerging technology, blockchain in (Zhang et al., 2021). Technology stage division is conducted using S-curve by the number of cumulative patents released each year. They identified three stages emerging stage in 2014-2015, slow growth in 2016-2017, and rapid growth after 2018. Then using individual datasets from these stages, technology topic analysis was conducted.

In their recent study, (Kim and Geum, 2021) proposed a data-driven roadmapping approach by integrating layer-mapping, contents mapping, and opportunity finding methods. They implemented their proposed method on self-driving (autonomous) vehicles. For the technology and market layers of the technology roadmap, they utilized individual data sources. For the technology layer, they utilized data from patent databases, and for the market layer, automobile magazines, community websites, and consumer reports. Content mapping comprises keyword network analysis for identifying major trends in each topic identified. Link prediction for opportunity finding, can be used as a substitute for technology workshops in traditional T-Plan based for assessing future technology prediction, the convergence of technologies, and changing market trends.

By integrating data analysis tools into technology roadmaps (Feng et al., 2022) aims to identify opportunities for metabolic disease drugs. Similarly, they follow layer mapping, content mapping, and opportunity finding. A topic model is created using BERT topic model on drug patent data; subject-action-object analysis is conducted for content mapping and link prediction for identifying potential connections, shown in the roadmap. This study uses deep learning and expert opinion for classifying documents as technical and market data, for mapping them to layers of roadmap.

(Miao et al., 2022) state that technology roadmapping is a practical tool for mapping emerging technologies to the market, and is traditionally done using expert opinions.

They use Technology-relationship-technology (TRT) analysis for identifying the relationship between products, functions, and technologies using smart wear patent data, and integrate it into the technology roadmap. Analyzing and discussing, future changes and trends in smart wear technology with the help of experts.

As mentioned earlier, there are situations when conducting a workshop or accessing to an expert is not feasible, such case in (Garza Ramos et al., 2022) where a roadmap for a start-up company is implemented in an emerging technology. The literature review method is used for identifying technology and product features of 3D Bioprinting and Cell culture technologies, also with a SWOT analysis. QFD and linking grids methods are used for linking market drivers and product features.

There are also hybrid methods, not relying only on quantitative methods but utilizing them and integrating them with traditional workshop-based qualitative technology roadmapping methods.

Similar to our study, (Zhang et al., 2016) used National Science Rewards scientific database search results from titles and abstracts on big data research, as their data source. They do a K-means clustering for finding clusters on big data research to form the components of the technology roadmap. They compared their clustering with LDA and Hierarchical Aggregative Clustering (HAC). They calculate TF-IDF similarity scores for forecasting how trends change in the future and identifying the relationships between the elements in the technology roadmap, combining quantitative analysis with expert knowledge.

(Li et al., 2015) integrates bibliometrics with expert opinions for constructing technology roadmaps on emerging technologies, applied a case study in dye-sensitized solar cells. (Li et al., 2019) combines text mining and expert opinions for analyzing technology evolution paths of perovskite solar cell technology, examining patents and scientific papers.

(Wang et al., 2018) analyzes the current situation and future development of nanogenerator technology trends in China, combining bibliometrics, patent analysis, and TRM. External factors and forecasts on future developments of trends are identified using expert opinion, and discussed in workshops.

(Ma et al., 2021) proposes another hybrid methodology containing topic modeling, SAO analysis, machine learning, and expert opinion for finding potential technological opportunities in dye-sensitized solar cell patents.

Technology roadmapping with quantitative methods increases reliability in decision-making and reduces expert bias, however, it may decrease the legitimacy of roadmapping in stochastic environments. (Ozcan et al., 2022) uses text mining in patents for clustering and creating a roadmap for technologies used in retail marketing. Apart from roadmaps, they separated the product layer as back-end and front-end. The proposed hybrid methodology is believed to minimize the reliability and validity issues of using a quantitative method alone.

Scenario analysis is used in various studies to make the roadmap more robust to different scenarios. (Noh et al., 2021) use scenario analysis and workshops for identifying potential products and services for 5G mobile services. They use patent analysis and TOPSIS to identify key technologies and capabilities, and use QFD for linking products/services to technologies, then finalize the roadmap with a workshop. (Jeong et al., 2021) tried to integrate risks into the roadmap as a layer. Creating a futuristic database by using websites, LDA is used for deriving possible future events and risks. Using co-occurrence and similarity of keywords from documents, a Bayesian network is created.

Table 1: Data-driven roadmapping methodologies and roadmap typologies

Study	Dataset	TRM Layers	Identify Roadmap Elements	Linking Roadmap Elements	Time	Temporality	Context
(Zhang et al., 2013)	Patents and Papers	Materials, Technologies, Products	Term Frequency Analysis, Experts	Association Rules, PCA	Years	Retrospective	Electric Vehicles
(Zhang et al., 2014)	Papers (WoS and EI Compendex)	Materials, Technologies, Products	Term Clumping, (TF-IDF, Clustering	TRIZ, SAO, Expert opinion	Years	Retrospective	Dye-Sensitized Solar Cells
(Jin et al., 2015)	Patents, Users Manuals	Technology, Product, Market	Text Mining, Patent Analysis	Cosine Similarity, QFD	Years	Retrospective	Solar LED Lightning
(Jeong and Yoon, 2015)	Patents	Patent Groups, Technologies	Text Mining, Patent Analysis	Cosine Similarity, SCAMPER scoring	Years	Prospective (7 years)	Amoled Display

Table 1 Continued:

(Geum et al., 2015)	Apple Website, Internet	Service- Product Mapping	Text Mining	Association Rule Mining, UCINET	No	-	Apple Products and Services
(Son and Lee, 2019)	Ministry Planning Reports	Material, Device, Soft ware	Expert Workshop	Relative Importance Fuzzy Inference, Clustering	Years	Prospective (6 years)	3D Printing
(Zhang et al., 2021)	Patents	None	Topic Modeling LDA	Cosine Similarity (Limitation: Should be combined with Experts.)	Growth Stages by Years	Retrospective	Blockchain
(Zhang et al., 2016)	Natural Science Foundation Awards Database (Scientific)	TF-IDF Scores	TF-IDF, LDA, HAC	Similarity Scores, Expert Opinion for classifying topics and locate on map	Years	Retrospective (Prospective After 2014)	Big Data Research

Table 1 Continued:

(Kim and Geum, 2021)	Patents for Technology, Magazines and Websites for Market	Technology, Market	LDA, evaluated by domain experts and mapped in TRM.	Link Prediction	No	-	Self-Driving Cars
(Feng et al., 2022)	Patents, Journals, Reports	Technology, Market	BERT and Expert Opinion, LDA, BERTopic, SAO	Link Prediction	Growth Stages by Years	Retrospective	Hyperuricemia Drugs
(Miao et al., 2022)	Patents	Technology, Function, Product, Market	TRT and Expert Knowledge ,TF-IDF, Similarity	Expert Knowledge based on TRT results	Years	Retrospective	Smart Wear
(Wang et al., 2018)	Patents, Papers	Technology, Industry, Production, Market	Bibliometrics, Patent Analysis, Expert opinions	Expert Workshops	Years	Prospective	Nanogenerators

Table 1 Continued:

(Ma et al., 2021)	Patents	Materials, Technology, Products, Market	LDA	Expert Opinion, SAO	No	-	Dye-Sensitized Solar Cells
(Ozcan et al., 2022)	Patents	Technology, Product/service/process Back-end, Front-end, Market	Text Mining, Expert Opinion	Expert Opinion	Growth Stages by Years	Prospective (7 years)	Retail Technologies
(Jeong et al., 2021)	Web Sites (Futuristic Database)	Risk, Technology, Product/System/Service, Market	LDA	Co-occurrence, Cosine Similarity, Bayesian Network (Roadmapping validated with patents and experts)	Years	Prospective (1 year)	Driver assistance systems
(Garza Ramos et al., 2022)	Papers and Expert Opinion	Business, Market, Product, Technology, Resources	SWOT, Literature Review	QFD, Linking Grids	Years	Prospective	3D cell culture workstation

2.9.3 Value Roadmapping for Emerging Technologies

There is a lack of financial and quantitative data for emerging technologies with uncertainties in the market such as edge computing; therefore, financial valuation and quantitative decision-making approaches cannot be utilized. In qualitative approaches where the opinions of decision makers are discussed, sometimes scored, and evaluated, (Dissel et al., 2009) discuss a lack of orientation in early-stage technologies when not enough expertise is present in the stakeholders and state many companies do not consider evaluation until the technology becomes more mature. Until then, they rely only on the expertise of managers which is sometimes referred to in the literature as the“gut feel”.

Therefore they propose value roadmapping for providing a useful environment for linking technological and business perspectives throughout the technology lifecycle (Cetindamar and Phaal, 2017; Dissel et al., 2009). The environment structures the individual expert opinions using a set of workshops for achieving the goals stated as steps of value roadmapping below.

Steps of value roadmapping are as follows:

- Define strategic framework, vision and scenario,
- Map technology development and investment milestones
- Define value streams
- Map market and business trends and drivers
- Map barriers and enablers
- Review project plan and value roadmapping
- Present visualization
- Maintain value roadmapping as a process

2.9.4 Data-Integrated Roadmapping

Roadmapping for data (Han and Geum, 2020) framework was proposed to integrate data as a layer of technology roadmap and the roadmapping process for increasing utilization of numerous big data produced by smart products and services. They underline digital transformation of services and products requires systematic planning focused around data. They propose three types of data integration: data as supporter, data as mediator, and data as value generator. Supporter data is for supporting the current processes without increasing the level of digitalization. As mediator, data is used for optimizing and increasing the efficiency of current processes, increasing the level of digitalization. Finally, as value generator, there is a feedback mechanism from smart services and products, resulting highest level of digitalization processes among data integration types.

A data science roadmapping framework is proposed (Kayabay et al., 2022) suggesting a tailored workshop-based methodology to overcome data, organization, technology and strategy (DOTS) related challenges for an organization to become data-driven. The framework integrates CRISP-DM process model thus considering the data lifecycle as a whole. The integrated data layer consists of data sources, data science processes, metadata objects, data products, and services sublayers in the roadmap architecture.

2.10 Literature summary and knowledge gaps

When edge computing literature is investigated, a lack of studies from the technology management perspective is observed. Many academic studies regarding adoption, maturity, strategic planning aspects, and other technology management activities are conducted on more mature technologies such as cloud computing. Those aspects should be also focused on edge architectures to establish a similar literature as the technology matures. However, edge architectures are generally use case dependent and not as generic as cloud solutions.

Considering the knowledge gaps identified in edge computing literature, technology management literature is investigated with a focus on emerging technologies. Technology management activities and tools are explained and alternatives are evaluated. Considering available data and similar studies focusing on emerging technologies with similar research questions, data-driven technology roadmapping studies are investigated in detail in Chapter 2.2.8. Table 1 presents a classification of studies based on their data source, methodologies used, and typology of the roadmap.

In their conclusion, (Kim and Geum, 2021) state a potential application can be a dynamic topic model for considering how trends are relatively changing over time and reflecting it in a data-driven technology roadmap.

(Kayabay et al., 2022) addresses the absence of a data-centric planning point of view for strategical roadmapping of an organization to become a data-driven organization, by identifying challenges they face and considering data as a separate layer in the roadmap. In the discussion for the future research section of the data science roadmapping framework study (Kayabay et al., 2022), it is stated that research technology assessment and forecasting tools can be integrated to complement the data-integrated roadmaps.

This thesis identifies the following knowledge gaps in the literature:

I. The literature lacks a study that applies a technology management framework to edge computing, neither organizational nor sectoral level.

II. No study in the data-driven roadmapping literature uses a dynamic topic model for creating the technology roadmap of an emerging technology.

III. Data-driven roadmapping studies for emerging digital technologies such as blockchain, cloud computing, or smart cities, do not consider data as a resource and a subject to be planned or investigated in the roadmap.

CHAPTER 3

RESEARCH METHOD

3.1 Research Questions, Aim, and Objectives

This thesis aims to fill the research gap mentioned in Chapter 1.2 on analyzing the sectoral edge computing adoption and implementation of emerging technologies into edge architectures. The research aim has been transformed into the following research questions:

RQ-1. What kind of real-life applications of edge computing and case studies are used in the IIoT environment?

RQ-2. What are edge computing trends and how do business objectives and market trends shape edge architectures regarding technologies used?

RQ-3. How to assess edge computing technology research using a data-centric approach?

Accordingly, these research objectives are formed:

RO-1. Investigating applications of edge architectures in industrial environments. Understanding motivations behind using these technologies.

RO-2. Identifying and mapping the relationship between technology and market trends of edge computing.

RO-3. Developing a data-integrated edge computing technology roadmap for IIoT.

3.2 Research Methodology

After defining research objectives from research questions, the methods selected for the objective and the data used in that step are identified in this section and shown in Figure 1. Corresponding to the first research objective, this research starts with a systematic literature review on edge computing applications in the industrial internet of things. The SLR reviews applications of cloud-to-edge architectures in the industry and groups them according to business objectives and application areas. It has been structured based on (Kitchenham, 2004). guidelines Snowballing approach is combined with a database search to systematically find case studies and real-life examples of edge computing applications in the industry regarding IIoT environments. Surveys and literature reviews on edge computing generally have a similar approach to a literature review, explained in Chapter 2.7 and Chapter 4.2.

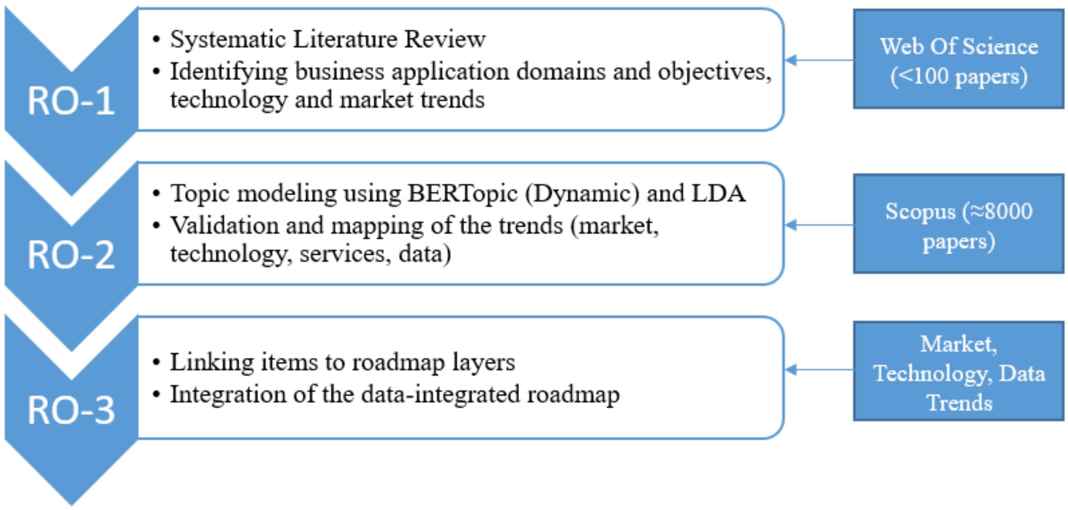

Figure 1- Research Objectives, Used Methodologies, and Data Sources

After the literature review on edge computing applications in IIoT, a knowledge gap in strategic aspects of applying new technologies to edge architectures in IIoT systems is identified. Although studies investigate edge computing on the technical developments side, they lack focus on business goals and which technology is used in which use case. One contribution of the SLR in Chapter 4 to the edge computing literature is that it investigates how different architectures are applied in a business objective, defining motivations of using an edge architecture along with the objective of the study itself.

Regarding the second research objective for identifying trends, some of the evident market and technology trends were identified in the SLR and presented at the end of Chapter 4 part followed by a discussion, that those identifications are made using the judgments of the researchers, hence those results may contain subjectivity. Therefore, following the SLR, a quantitative approach, unsupervised learning by topic modeling in a broader database search is used, without the inclusion criteria. Data for this section consists of a database search of titles, keywords, and abstracts of academic papers. Identifying the same trends in the same topics, the results of SLR are validated and expanded in content. Supporting SLRs using topic models or other text-mining models can be observed in other research domains (Asmussen and Møller, 2019), reducing researcher bias, and increasing reproducibility. (Ho and Lee, 2022) uses Human-in-the-loop topic modeling for identifying their roadmap structure, which is similar to our approach, provided with both researcher expertise and insights from SLR results.

LDA and BERT topic models are trained, using different numbers of topics, and smaller different versions of the dataset and several experiments on preprocessing methods such as lemmatization have been conducted and tracked using text files. When studies that embrace similar data-driven approaches to form a technology roadmap for emerging technologies are investigated in Section 2.2.9, most of them are published in Technological Forecasting & Social Change journal (Kim and Geum, 2021), (Zhang et al., 2021), (Garza Ramos et al., 2022), (Li et al., 2015), (Li et al., 2019), (Ma et al., 2021), (Noh et al., 2021), (Zhang et al., 2016). We used LDA and BERTopic model results to identify the market and technology trends and the relationships between them to be visualized as elements in the technology roadmap.

They both provided reasonable results that validate SLR results and both results produced interpretable value by the researcher, so one could not be chosen as superior to the other. Dynamic topic modeling using the BERTopic library (Grootendorst, 2022) enables one to trace and visualize how topics frequency changes over time, addressing the potential future research identified by (Kim and Geum, 2021) on being able to trace how trends change by time, explained at the end of Chapter 2. Market, technology, and data trends are identified using keywords appearing in the same topics (clusters), explained in detail in Chapter 5. In general, LDA with a lower number of topics was consistent with SLR results, and BERT with a slightly more number of topics provided good insights for the identification and mapping of the trends.

After identifying market and technology trends, this study presents data sources and applications used in applications from the edge computing applications literature, detailing topic model findings. All the identified elements are then integrated as different layers of a technology roadmap using expert opinion, SLR, and topic modeling findings, as shown in Figure 2. To provide ease of tracking to the readers, little circles next to process rectangles indicate the chapter number in which the process is located in the thesis.

In this study, the nature of research questions 2 and 3 needs investigation of very different edge applications rather than focusing on a specific case, unlike industrial applications in RQ-1. Therefore, it is a more breadth-first exploratory investigation rather than depth-first (See 5.2 data collection for topic models). A study on how companies can develop a digital strategy (Al-Ali and Phaal, 2019) also used both qualitative and quantitative methods to answer different research questions. By customizing technology roadmapping for digital transformation, they conducted case studies and applied workshops in organizations. To identify digital strategy archetypes, (Al-Ali et al., 2020) utilize text classification on Fortune 500 Earning calls, doing cluster analysis with the best performed model, a pre-trained RoBERTa.

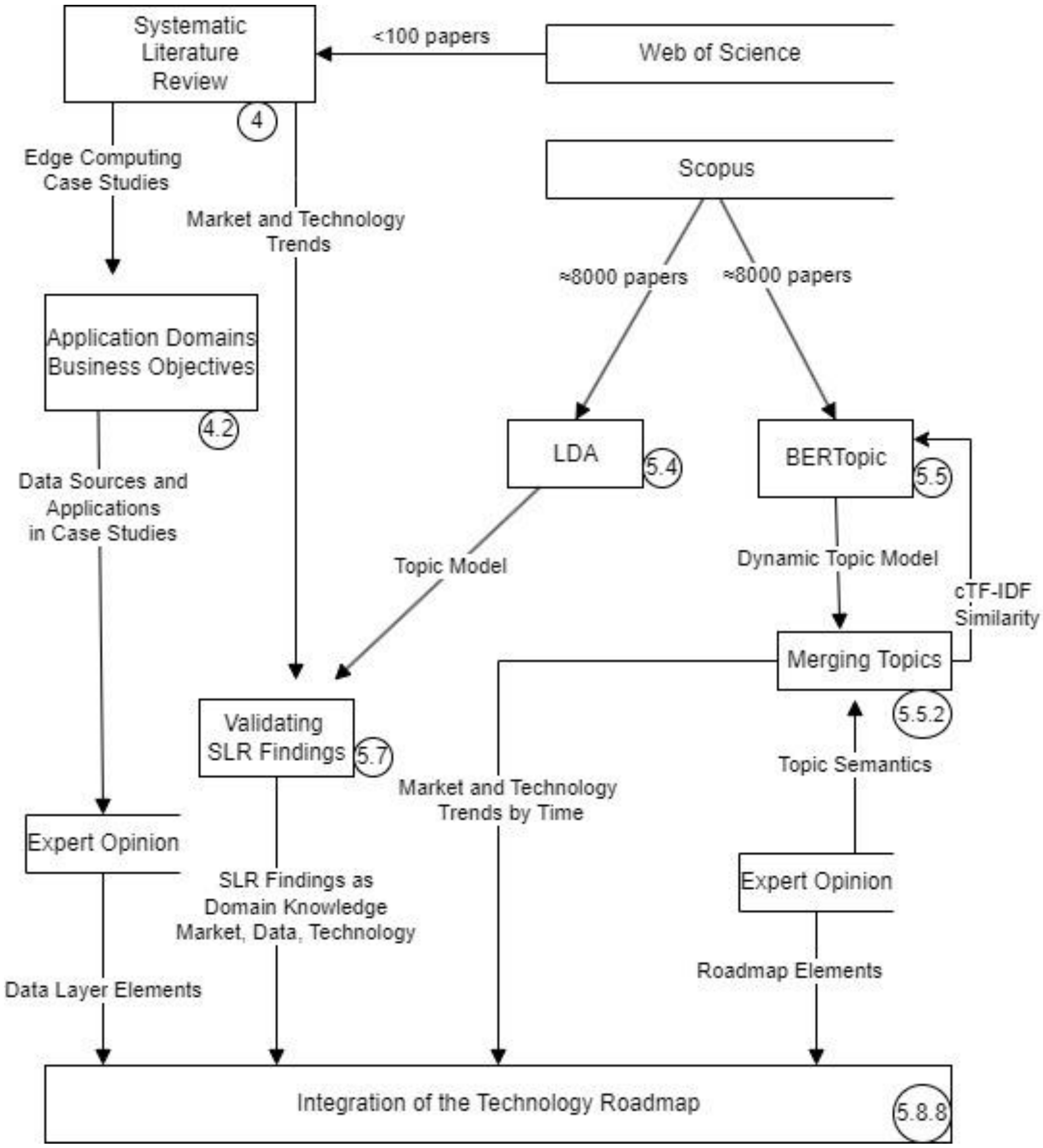

Figure 2- Research Methodology Data Flow

CHAPTER 4

EDGE COMPUTING APPLICATIONS IN INDUSTRIAL IOT SYSTEMATIC LITERATURE REVIEW

4.1 Systematic Literature Review Methodology

This literature review is conducted based on Kitchenham's guidelines (Kitchenham, 2004; Kitchenham and Charters, 2007). Snowballing approach is combined with a database search to systematically find case studies and real-life examples of edge computing applications in the industry regarding IIoT environments. Accordingly, the research questions of the study are formed as follows:

- What kind of real-life applications of edge computing and case studies are used in the industry regarding an IIoT environment?
- What are the main objectives for investigating edge computing architectures?
- How edge architectures are shaped regarding these objectives and nature of the business problem?

In this thesis, applications of edge computing in the industry are investigated and grouped according to their objectives and application areas. The aim is to present the current implementation status of edge technologies and trends to assist organizations in researching their business objectives.

4.1.1 Search Process

The following search query is used as a starting point of the search: ((“Edge”OR “Fog”) AND “Computing” AND “Industrial” AND “IoT”). First, the Web of Science is used as the main database and results are enriched using IEEE Explore in snowballing phase, too. Search results are stored in MS Excel. 67 initial results were found after searching the Web of Science Core Collection. Only four of them were released before 2017, and the number of published papers had increased by the year 2021. 15 studies were selected after the initial elimination. After fully reading these articles, and including references of the references of those studies, 28 papers were selected as **Primary Studies** (PS). The studies focusing on edge computing use cases, applications, and architectures in IIoT, covering algorithms, tools, benefits, and implementation challenges were regarded as PS, and 16 of them are journal articles, while 11 studies are conference papers, and one is a book chapter. The authors conducted iterative meetings for identifying the main objectives and trends for grouping papers.

4.1.2 Inclusion Criteria

Studies on edge computing use cases and applications in IIoT, focusing on algorithms, tools, benefits, and implementation challenges such as energy consumption or computing capabilities of different architectures in real-life scenarios, were included. Domain specific IoT articles with less or no edge computing implementation were excluded from the search. Studies targeting smart cities, blockchain-edge, autonomous-driving, 5 and healthcare domains rather than manufacturing were excluded. The papers that include only technical solutions in computation load balancing, energy consumption, security, and data storage efficiency were included if they explain a particular case study application related to the IIoT domain. In the snowballing phase, papers are found with English titles and abstracts, but Turkish content is also included.

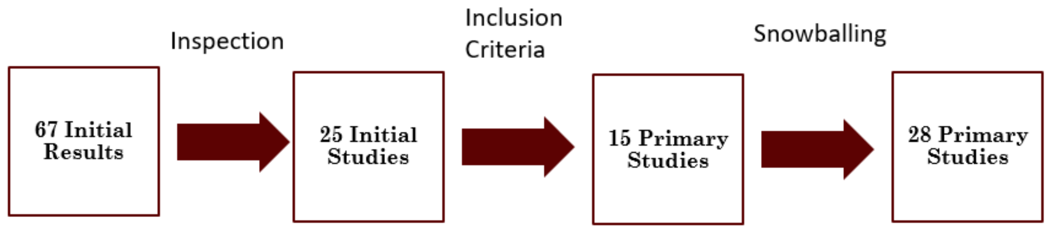

Figure 3- Literature Review Process

4.2 Systematic Literature Review Findings

This section reviews applications of edge architectures in the industry and groups them according to objective and application areas. Studies are explained regarding how the architectures and trends are shaped to fulfill the objectives, allowing practitioners to have an idea while developing implementations for their cases. Studies focusing on security, latency, resource utilization, and energy efficiency are presented in Table 2.

4.2.1 Open-Source Edge Architectures

An ***open-source*** architecture for industrial networks called IFog4.0 is proposed in (Ghazi Vakili et al., 2019) with case studies in an emulated gas regulation station environment. A Fog-Management module has been developed to manage Docker containers. Only this component is exposed to the external internet; Docker's internal security and a firewall limits access to this network. Using Node-RED programming working with a flow-based programming paradigm, nodes of inputs and outputs developed with NodeJS, communicate in a network. The Grafana environment provides data visualization, using the data provided by PLC Siemens through the s7 Profinet protocol. Odoo ERP, MySQL, and MongoDB databases have been used for storage. For Linux Kernel to work with Profinet, Modbus, OPC UA, industrial communication drivers were implemented. The author states their goal was to provide this technology for SME's digital transformation.

Although there are similar layered architectures of edge computing, the one proposed by (Chalapathi et al., 2021) is specifically for manufacturing and is presented with a case study in the paper. Application domain provides monitoring and control services, data domain provides data cleaning, feature extraction, and operation throughput optimization, increasing system efficiency. Network domain manages devices using ***SDN***. Time Sensitive Network (TSN) protocol is also employed to make the architecture time-sensitive. Finally, device layers consist of physical infrastructure, sensors, and actuators, which must be flexibly sustainable in an environment with various protocols, and dynamically changing system execution strategies depending on data. A similar architecture is used to compare the productivity of a newly deployed edge architecture to an existing private cloud on a candy packaging production line (B. Chen et al., 2018). Standardized Data Distribution Service (DDS) middleware protocol and Ethernet were integrated into the network between edge devices. Tasks were individually assigned to each robot which the system can shift in case of any failure using Contract Net Protocol (CNP). Robots can bargain and make agreements on their own to complete the assigned task. The Hadoop architecture was built at the local database level to set up this system, and real-time analysis was claimed to be performed with MapReduce. Machine status and sensor data for the model were loaded from a Raspberry Pi. An OPC UA server was operating to transmit raw data from sensors and handle preprocessing tasks. Although the network's speed decreased from 16MB/s to 6 MB/s after switching to the edge, results show in high-volume mass production, edge provides more productivity.

A manufacturing process control system has been proposed by (Wu et al., 2017) for monitoring production lines, collecting and analyzing data to increase efficiency. First, stream data collected from sensors through communication adapters working on OPC UA and MTConnect protocols. Then, an edge node, streaming data in real-time, provides control signals allowing the system to work in low latency. Samples from data can be transmitted to cloud centers for further analytics, and building models. An ***open-source edge architecture*** is implemented in a university campus's air quality monitoring system (Kristiani et al., 2021). Data was collected from Arduino sensors across the campus. Low Power Wide Area Network (LoRA) gets data from sensors. It transmits data to the LoRa gateway, then to an edge gateway of Kubernetes minion installed on a Raspberry Pi for final unified delivery to the data center. MQTT protocol enables sending alerts to devices if there are anomalies in the data. The network's overall performance is analyzed using Ganglia Monitoring System, and comparisons are made to achieve the best open-source architecture.

4.2.2 Security Applications

There are many studies related to distributed ***blockchain*** and edge computing convergence in the literature related to the security of IIoT systems, such as (Kumar et al., 2020) (Lee et al., 2020; Wu et al., 2021). An approach that uses blockchain and context-aware security for IIoT environments was proposed by (Portal et al., 2020), and implementation in an additive manufacturing site was presented. Data for context reasoning is provided with blockchain instead of the cloud, claimed to be providing a better fit for edge devices, reducing communication costs, and increasing bandwidth efficiency.

Another tiered edge architecture, (Sittón-Candanedo et al., 2019) utilized blockchain to data from sensors for increasing security in the Edge layer. The proposed architecture is implemented in an agroindustry platform to monitor and support decisions in a dairy farm, implementing AI to detect anomalies. Sensors for measuring rain, wind speed, ground temperature, and humidity, gas sensors for detecting chemical levels in the air, and biometric sensors for measuring body temperature, heart rate, and breathing rates for cows were used. ZigBee communication was used to connect IoT gateways and sensors. Edge layer node is a Raspberry Pi for preprocessing IoT data and forwarding it to the cloud. A data processing framework is proposed in (Fu et al., 2018), enabling secure data storage and operations using edge in IIoT systems. Data management and encryption challenges are summarized, and solutions were proposed as a framework. Then it is evaluated using simulations and experiments in a prototype of a system built to monitor the temperature in a factory. Another real-time industry application for edge computing was implemented in a simulated smart factory to observe the effectiveness under cyber security scenarios (Güven and Çamurcu, 2018).

4.2.3 Energy Efficiency and Resource Utilization Applications

Energy efficiency and computational workload are vital aspects to consider while utilizing edge nodes. SDN is widely used for managing a network of middleware devices, reviewed by (Kaur et al., 2018), where the trade-off between energy efficiency and latency is evaluated. Djemame proposes a technical architecture that leverages SDN with Network Function Virtualization (NFV) and serverless architectures to reduce the high energy consumption of edge architectures (Djemame, 2021).

An adaptive data transmission algorithm using SDN and edge computing for IIoT is proposed by (X. Li et al., 2018) to find an optimal route for traffic load, task deadlines, and energy consumption. To achieve an industrial internet, Chen et al. (C.-H. Chen et al., 2018) proposed a framework consisting of high-level embedded microcontrollers and gateway systems to provide an edge gateway of a smart sensor Fieldbus network. With the help of distributed computing, the gateway efficiently performs network management, data collection, and communication, considering power consumption and providing better scalability than traditional IIoT solutions. A similar problem is modeled using a probabilistic approach (Chekired et al., 2018). A priority queuing model is implemented for scheduling IIoT data according to priorities as high and low, formulated as a mixed non-linear integer program, and an optimal solution is found using branch and bound algorithm with simulated annealing. Also, an IIoT offloading algorithm for queuing and processing work according to resource priority and availability is proposed.

Utilizing ***deep learning*** requires high computation power and bandwidth; the cloud requires data transfer while the edge requires expensive computing resources. In an edge architecture tailored for deep learning (Liang et al., 2020), complexity is optimized in line with the computational capacity of edge devices. To evaluate the solution, the authors formed a convolutional neural network using real-world IIoT data and apply their approach, doing experiments that reduce network traffic while maintaining the model's classification accuracy. In the scenario, they worked with 30 different components used and identified by cameras, representing a production line. One way of processing deep learning in MEC is by inferencing, and executing pre-trained models with newly generated visual content from mobile edge devices. The study by (Xu et al., 2021) formulated the inference offloading problem to minimize energy consumption and evaluated the performance of proposed algorithms using simulations. In order to use deep learning for anomaly detection, (Ferrari et al., 2019) performance of different architectures are tested. Trade-offs of choosing the architecture considering scalability, bandwidth, and delay have been presented. The author concludes that scaling cloud computation power results in full cloud outperforming the edge. To monitor the real-time status of machinery and conduct predictive maintenance, a database is created by (Oyekanlu, 2017), small enough to fit in memories of edge devices using open-source Python SQLite. The data is processed in edge devices, analyzed, and only the recommendation is sent to the central cloud to take action.

Containerized edge architectures are evaluated by (Liu et al., 2021) in terms of industrial requirements, measuring round trip time, bandwidth, processing capabilities, and latency while doing machine learning tasks for predictive maintenance. Microsoft Azure IoT Edge is utilized for running container applications on Raspberry Pi. Device-to-cloud, device-to-edge-to-cloud, and device-to-edge architectures are compared. Results showed that containerization does not decrease performance while increasing flexibility and scalability. Fog Computing Platform reference architecture is proposed for IIoT applications by (Pop et al., 2021), using open standards, OPC UA, and TSN. In the Conveyor Distribution System, a machine is provided packages containing tags, and the system delivers them to the destination by accessing a database by reading the tag. Electric motors that provide the movement, forming the "machine level" of architecture, produce massive data, which is problematic to send to the central cloud. Network configurations, network traffic, and other benefits of using an edge architecture are explained in detail in their paper about this use case (Barzegaran et al., 2020). The authors also mention different use cases of the same architecture, such as industrial robotics on the shop floor and machine control using edge platforms. These cases are also referred to as Fog-based Industrial Robotic System and next generation of machine control using a Fog Platform in (Shaik et al., 2020) and (Denzler et al., 2020).

Table 2: Papers are grouped by their objective and application

Papers	**Main Objective**	**Application Domain**
(Sittón-Candanedo et al., 2019)	Security	Agricultural Monitoring
(Portal et al., 2020)	Security	Additive Manufacturing
(Güven and Çamurcu, 2018)	Security	Simulated Factory
(Ghazi Vakili et al., 2019)	Resource Utilization	Real-Time Gas Pressure Control
(Oyekanlu, 2017)	Resource Utilization	Predictive Maintenance
(Chalapathi et al., 2021)	Latency + Energy Efficiency	Active Maintenance
(Kaur et al., 2018)	Latency+ Energy Efficiency	Software Defined Network

Table 2 Continued:

(Kristiani et al., 2021)	Energy Efficiency+ Resource Utilization	Air Quality Monitoring
(C.-H. Chen et al., 2018)	Energy Efficiency+ Resource Utilization	Smart Manufacturing System
(B. Chen et al., 2018)	Security+ Latency	Active Maintenance
(Okay and Ozdemir, 2016)	Security+ Latency	Smart Grid
(Kumar et al., 2020)	Security+ Latency	Simulation
(Fu et al., 2018)	Security+ Resource Utilization	Factory Temperature Monitoring
(Pop et al., 2021)	Latency+ Security+ Resource Utilization	Conveyor Routing, Distributed Predictive Maintenance
(Chekired et al., 2018)	Latency+ Resource Utilization	Simulation with real IIoT Data
(Denzler et al., 2020)	Security+ Resource Utilization	Real-Time Machine Data Analytics
(Liang et al., 2020)	Latency+ Resource Utilization	Image Classification Simulation
(Ferrari et al., 2019)	Latency+ Resource Utilization	Real-Time Anomaly Detection
(Lee et al., 2020)	Latency+ Resource Utilization	Numerical Experiments
(Liu et al., 2021)	Latency+ Resource Utilization	Vertical Plant Wall System
(Shaik et al., 2020)	Latency+ Resource Utilization	Industrial Robotics
(X. Li et al., 2018)	Latency+ Resource Utilization + Energy Efficiency	Smart Manufacturing

4.3 Discussion

Although there is research on when and why to use edge architectures, all applications are experimental. It is found that there is a lack of a strategic deployment decision framework for the implementation of such technologies. All applications are trial-based, without sufficient information about organizational research and feasibility, which is generally the case in a large company. Working with academia and companies with expertise is vital for understanding the best practices. The research highlights:

- Real-life applications of edge computing do not consider the market trends and adoption strategies mentioned in Section 2. There is a severe problem in choosing the right technological solution among an abundance of alternatives for decision-makers, application developers, domain experts, and operational personnel (Gokalp et al., 2016).

- Latency is a common concern among edge applications in IIoT. Depending on the complexity of the computation business case requires, e.g., a deep learning task or real-time monitoring, literature presents offloading mechanisms or tailored approaches for efficiently using resources.

- Following trends of edge computing are identified. First, open-source tools have been commonly used to develop edge architectures. Second, blockchain has been incorporated into IIoT networks for security. Third, approaches such as SDN, containerization, and computational offloading algorithms are used for resource utilization and energy efficiency.

- Edge computing's objectives and trends coincide with the IT/OT convergence trend.

4.4 Limitations

Limitations of the SLR are identified as follows:

- Snowballing and manually investigating specific titles may have reduced the reproducibility of the search process.

- For presenting trends, case studies, and discussion grey literature have been utilized including reports from non-governmental organizations, market research/consulting companies, and researcher’s experience.

- As stated in 2.3.1, groupings of papers by their objectives and identified trends resulted in iterative meetings between researchers may contain subjectivity.

CHAPTER 5

DEVELOPMENT OF A DATA-INTEGRATED EDGE COMPUTING TECHNOLOGY ROADMAP

5.1 Introduction

In an emerging technology research domain, scientific literature progresses rapidly. Chapter 4 briefly discusses the technological and business trends in applications. However, the generalizability of the findings and discussion are stated as a limitation, and also, prone to researcher bias. This chapter collects more scientific resources on edge computing applications in the industry from the academic literature and a more data-driven approach is used. Some of the findings of the systematic literature review are validated and business and technology trends are matched using dynamic topic modeling, also allowing to make predictions on how trends are changing with time. In contrast to using expert opinion-based qualitative approaches, using data-driven analytical methods would be less prone to human bias (Miao et al., 2022). Also, applying analytical methods such as text mining, topic modeling, and patent analysis enable the investigation of more data in a shorter amount of time, compared to manual approaches with the same goals (Nazarenko et al., 2022).

5.2 Data Collection

The search query is searched in the titles, abstracts, and keywords both for being able to identify the most relevant search results and also to avoid papers that include the keywords only in a small part of the text. The search query used in the Scopus database is similar to the one in Chapter 4 and is as follows:

TITLE-ABS-KEY ("edge computing" OR "fog computing) AND (industrial OR manufacturing) AND (iot OR "internet of things") AND (application OR applications OR "case study")

7773 documents were analyzed, and almost 60% were classified as articles followed by 31% as conference papers and 4.5% as book chapters. The number of documents per year is shown in Figure 4. However, with the effect of the Covid-19 pandemic, some conference proceedings were canceled and researchers might choose not to send their work to virtual conferences. To address this issue, Figure 5 shows the database search results when conference work is excluded. So we can say the increasing trend of the number of documents is preserved.

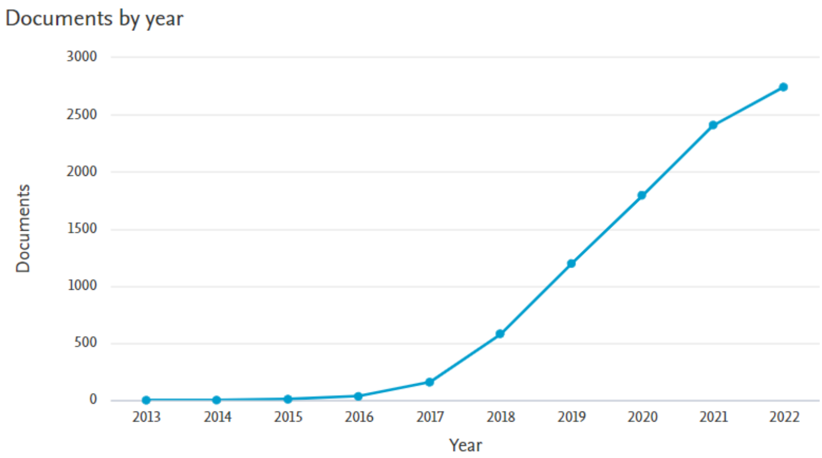

Figure 4- Search Results, Number of Publications by Years, from Scopus

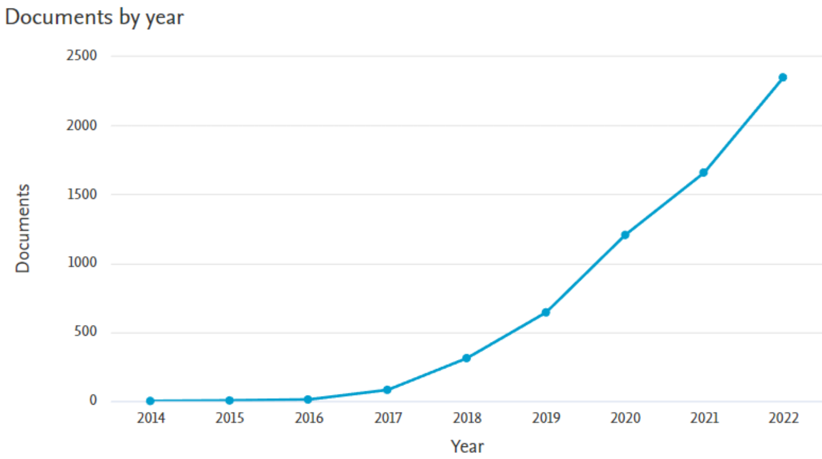

Figure 5- Number of documents by year after excluding conferences

In the SLR in Chapter 4, industrial environments are questioned as a sub-domain of using edge architectures and also for being able to manually investigate studies using a systematic literature review and developing expertise on the state of research. In this chapter, industrial or manufacturing keywords are again in the search string, however, there are also papers on autonomous vehicles, healthcare systems, and pollution monitoring systems, which were excluded from the analysis in Chapter 4 and are not excluded from topic modeling results. When the search results and topic models are investigated in terms of using edge technology, it can be seen that innovations and changes driven by novel applications in one domain affect the other. Strategically, big industrial organizations can choose to diversify or even change their target market such as Philips, which used to be more end-user electronics, now focuses more on the healthcare market (Balasubramanian, 2022). Therefore, with lots of studies focusing on diverse applications, we use broader and more holistic database search results to explore using topic modeling, compared to the manual systematic literature review.

5.3 Topic Modeling for Emerging Technology Analysis

Topic modeling is an unsupervised learning task that is often classified as a Natural Language Processing (NLP) application, where related terms are clustered under topics from a large set of written documents. In this chapter, studies that use topic modeling for similar objectives as our study are presented. A well-established state-of-the-art topic modeling algorithm, Latent Dirichlet Allocation (LDA) first applied in ML by (Blei et al., 2003) has been widely used in the literature, as presented in the examples below and detailed in Chapter 5.4.

Recent studies applied LDA in a variety of research areas. In this chapter, similar related examples are given corresponding to our research objectives stated in Chapter 3.1 for detecting emerging technology trends. One example study that applies topic modeling with LDA is (Ma et al., 2021), which proposes a hybrid methodology containing topic modeling, Subject-Action-Object (SAO) analysis, machine learning, and expert opinion methodologies, for finding potential technological opportunities in dye-sensitized solar cell patents. Similar to this study, they conclude and present their findings with a technology roadmap.

(Atzeni et al., 2022) used topic modeling with BERT for identifying convergence of machine learning for Wi-Fi connection analysis. They counted the number of occurrences of machine learning model names in each identified topic to find widely used ML methods for Wi-Fi data. (Inaam ul Haq et al., 2022) applied LDA topic modeling to Scopus abstracts and titles to detect IoT research trends over time, and also classified research orientation for different sectors where IoT is used.

(Kwon et al., 2022) uses LDA topic modeling in patents and for detecting technology trends, TF-IDF scoring is used which is explained in detail under Chapter 5.5.1, for clustering unstructured text data. The popularity of the trends over time is classified as hot, active, and cold research topics. The Generative Topographic Mapping algorithm is used for identifying vacant technologies. A similar method to this study is embraced by (Kukushkin et al., 2022), for visualizing and analyzing the bibliometric situation in digital twin research including 8693 abstracts of publications between 1993 and 2022. Topics are modeled using BERTopic and LDA. It is noticed that the number of studies in the search results is almost the same as for our search results for edge computing between 2015 and 2022, which can be a reference for increasing popularity.

To identify digital strategy archetypes, (Al-Ali et al., 2020) utilize text classification on Fortune 500 Earning calls, doing cluster analysis with the best performed model, a pre-trained RoBERTa. Each cluster represents the digital capabilities of companies are called a topic.

5.4 Latent Dirichlet Allocation

LDA is a generative model to discover underlying semantic knowledge as topics in a corpus of text data. Corpus in NLP refers to a collection of text data to provide a sample of language regarding patterns and structures to train the model. As can be seen in Chapter 5.3, LDA has been widely used in topic modeling tasks.

There are also latent semantic analysis and latent semantic indexing methods, which use singular value decomposition to reduce the dimensionality rather than a generative approach, also used in some studies, but have been largely replaced by LDA (Inaam ul Haq et al., 2022). The initial paper where LDA is proposed (Blei et al., 2003) compares it with probabilistic LSI (pLSI) and explains how LDA overcomes some overfitting problems in pLSI.

To determine the number of topics, coherence score, and human interpretability were evaluated by the researcher in a Human-in-the-loop manner while changing the number of topics hyperparameter. Coherence score is generally used for evaluating the number of topics hyperparameter. It measures the similarity between words within the same topic. Perplexity is also another evaluation method used in NLP, but some argue it does not correlate with human judgment (Chang et al., 2009). Therefore, in the experiments, there are cases where some evaluation metrics are worse but the topic outputs of the model are more coherent when evaluated by the researcher. Comparing LDA models, these metrics are calculated and represented in the results section to maintain reproducibility and objectivity.

5.4.1 Data Pre-Processing

There are several pre-processing steps applied to data before proceeding to LDA. Gensim and Spacy libraries of Python are used for pre-processing. NLTK library is used for removing stop words.

1. Each abstract had a copyright sign and a piece of publishment information at the end, these are also stripped from the text along with stop words. Text is converted to lowercase.

2. Words are tokenized using the Gensim library. Tokenization is a fundamental step in NLP tasks. It is splitting each word, sentence, or phrase in the text into small units called tokens.

3. Word phrases such as: “edge”+ “computing”, an example of a bi-gram, are detected and added to the corpus, as suggested by (Mikolov et al., 2013).

4. Lemmatization is applied to tokens to take roots of the words that semantically have the same meaning. It removes suffixes and returns the word to its dictionary form.

5. Cleaning the corpus from tokens that have too high or low frequency i.e extreme words. Filtering tokens appearing in more than 90% of documents would eliminate terms in the database search explained in Chapter 5.2 Another filtering was to eliminate tokens that only appear in less than 2 documents.

As discussed above on the evaluation metrics, both lemmatized and non-lemmatized models ran. Results and opinions on using lemmatization and filtering extreme tokens for topic modeling are presented in the following chapters. Experiment tracking has been conducted using Microsoft Excel and notebooks.

5.4.2 *LDA Hyperparameter Tuning*

The number of topics parameter needs to be optimized by the researcher. In the experiments, several model versions are tracked, applying different pre-processing steps of lemmatization, using bigram models, and different datasets where different levels of extreme tokens are removed from the corpus. Below, charts reflecting how the number of topics and the coherent score (c_v) (Syed and Spruit, 2017) change with the number of topics are given in Figures 6 and 7 below. Both indicate K=10 is a good number to use as the number of topics.

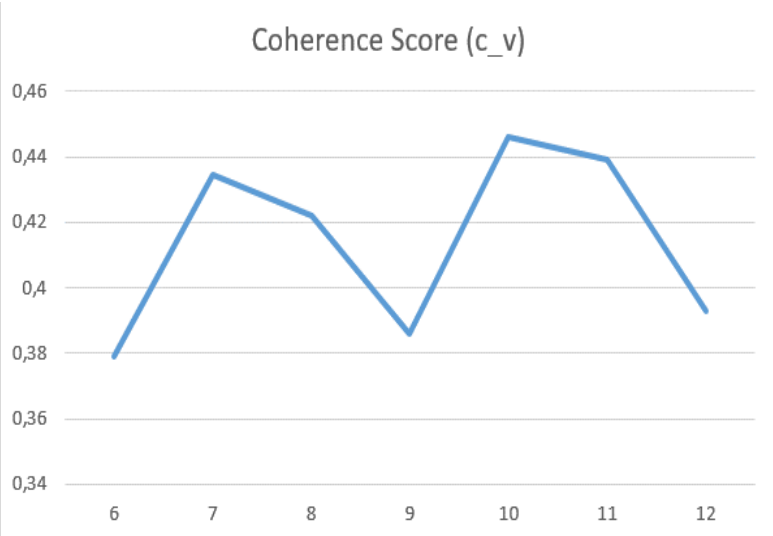

Figure 6- Finding the optimal number of topics in non-lemmatized model

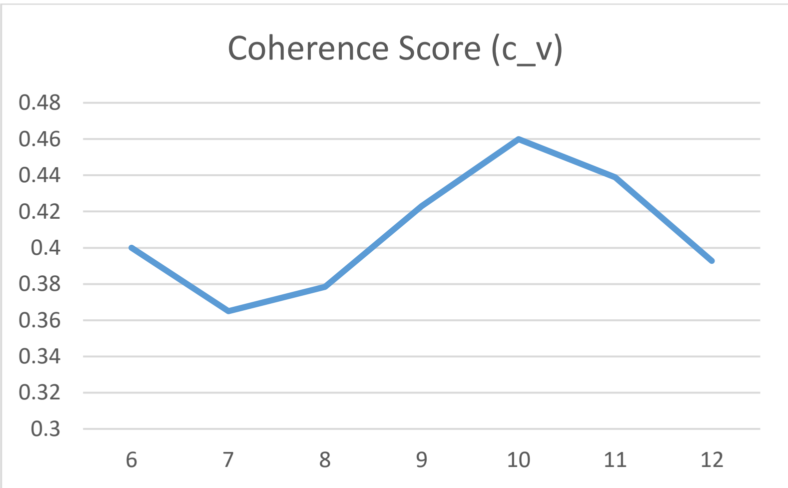

Figure 7- Finding the optimal number of topics in lemmatized model

In the experiments, perplexity scores of the non-lemmatized model performed better than the lemmatized versions with the same hyperparameters. However, as shown in Figures 6 and 7, it can be seen that lemmatization increases the coherence score. In the results, some of the findings in the non-lemmatized model are not found in the lemmatized model. Therefore, both models are utilized in terms of topics produced.

There are another two hyperparameters of α and β. In the experimentations, different alpha and beta values are tried in the model in two validation sets. Although some results were achieving higher coherence and perplexity scores, the interpretability of these models was poor. Our experiments and optimization resulted similarly to (Chang et al., 2009), optimizing hyperparameters using perplexity as the objective resulted in non-human-interpretable results with almost no insights. Also, (Hoyle et al., 2021) states that better coherence score in topic models are not always ending up with a preferred topic model therefore expert judgment is essential for evaluating and discussing the results of LDA. Similarly, in this study, the number of topics, alpha and beta parameters are tried to optimize using manual grid search and Optuna. The models with the best perplexity score or coherence scores did not result in the topic models with insightful clustering of topics (trends of market and technology) in a human-interpretable manner. Therefore, the 'auto' option in the Gensim library is used for the alpha and beta parameters of LDA (Hoffman et al., 2010) and different numbers of topics are compared. Therefore, they learned from asymmetric prior data, by setting 'auto' for these values in the Gensim library for Python (Hoffman et al., 2010), which gave the most interpretable results.

5.4.3 LDA Results

The final topics identified and the keywords in each topic are given in Table 3 below for the LDA topic model without lemmatization. See, in the results, "task" and "tasks", "offloaded" and "offload can be identified in the same topic because of not using lemmatization.

Table 3: Topics and keywords of the non-lemmatized LDA model

Topics	**Keywords**
Topic 1- Latency	Fog, cloud, devices, services, latency_ end, network, layer, low_latency, bandwidth, quality, available, qos, sdn, data_centers
Topic 2- Technology Forecasts	Technologies, systems, management, integration, hardware, big_data, future, platforms, networking, smart_cities, towards, connectivity, gateways, enabling, flexible
Topic 3- Resource Management	Edge_nodes, model, improve, mechanism, collaborative, business, content, realize, operational, cache, terminal, agent, grid, pricing, network_congestion
Topic 4- Energy	Algorithm, energy, energy_consumption, optimization, transmission, routing, scheduling, energy_efficient, consumption, clustering
Topic 5- Security	Data, privacy, secure, blockchain, protocol, attacks, image, sensitive, trust, authentication, sensor_nodes, data_aggregation
Topic 6- Computation Offloading	Task, offloading, (offloaded, offload), task_scheduling, path, parallel, battery, long_term, capture, transform, simulations
Topic 7- Smart Manufacturing	Industrial_internet, detection, deep_learning, production, accuracy, classification, equipment, machines, semantic, products, factory, recognition, smart_factory, time_series, robots, technical
Topic 8- Software Applications	Software, urban, tracking, city, low_power, events, open_source, inference, controller, web, project, health_monitoring, predictive, measurement, traffic_congestion, verification
Topic 9- Monitoring	Monitoring, sensor, healthcare, health, people, machine, modeling, water, patients, diagnosis, personal, data_acquisition, wearable_devices,
Topic 10- IoV	MEC, vehicles (vehicular), task_offloading, resource_allocation, radio_access, cellular, mec_server, reinforcement_learning, parking, base_station, joint_optimization

pyLDAvis library is used for creating the intertopic distance maps below (Sievert and Shirley, 2014). Also, keywords in topics are ordered using the relevance metric (Sievert and Shirley, 2014). In the intertopic distance maps, a larger topic circle indicates the higher frequency of the topic in documents. Distances between the topics reflect the distance measures used and can be interpreted as a similarity or divergence relationship between the topics. This visualization uses multidimensional scaling to project the distance between topics onto two dimensions by placing centers of circles (Chuang et al., 2012). Jensen-Shannon divergence is used as the distance metric and PCA is used for dimensionality reduction for the visualization by default pyLDAvis (Sievert and Shirley, 2014).

From the intersecting sets shown in Figure 8, we can understand there are common tokens in the topics. Topics 1 and 2 share common tokens such as: "applications", "architecture" and "requirements", which are not trend dependent but words occurring a lot. That's why they are intertwined in the intertopic distance map. It is noticed that topics 6 and 10 share common tokens such as "task_offloading". Similarly, topics 1 and 3 shares "resources", "network" and "performance" tokens.

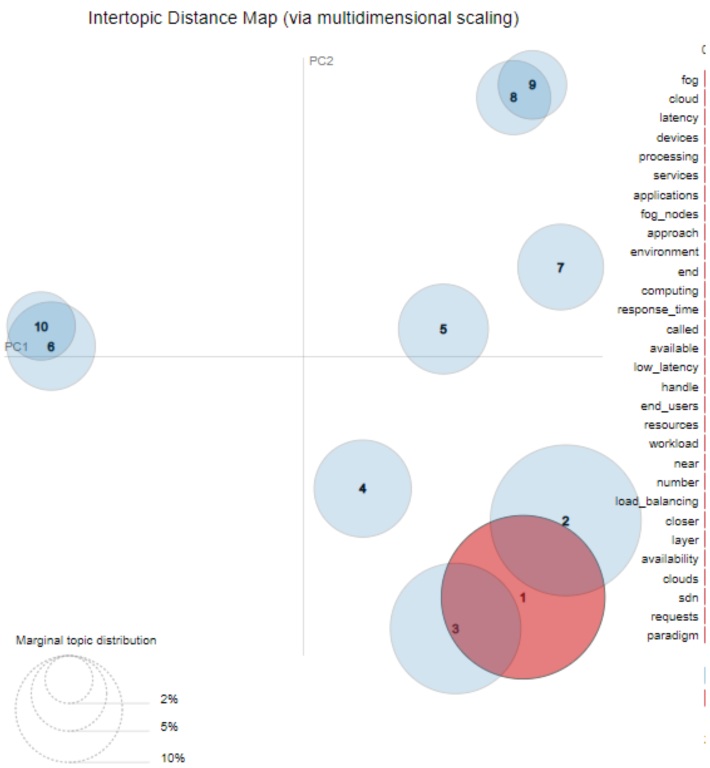

Figure 8- Intertopic Distance Map of Non-Lemmatized LDA

“Computation_offloading”, “network” and “optimal, tokens are shared between topics 6 and 10. This indicates computational offloading related research on the internet of vehicles, or autonomous driving market trend is conducted by researchers. Below, table 4 shows LDA model trained using the lemmatized tokens, by grouping different forms of the same word.

Table 4: Topics and keywords of the lemmatized LDA model

Topics	**Keywords**
Topic 1- Latency	Device, cloud, environment, approach, processing, latency, distribute, reduce, solution, performance, increase
Topic 2- IoV	Network, base, communication, vehicle, traffic, mechanism, dynamic, function, cache
Topic 3- Smart City	Technology, smart, management, platform, development, research, concept, discuss, integration, field, software, big_data smart_grid, smart_city
Topic 4- Edge Efficiency	Problem, energy_consumption, algorithm, delay, solve, optimization, cost, strategy, transmission, scheduling, simulation, minimize, resource_allocation,
Topic 5- Security	Security, scheme, privacy, secure, protocol, attack, blockchain, message, trust, encryption, sensitive, protect, authentication, decentralize, association, interface
Topic 6- Monitoring	Sensor, analysis, real_time, monitor, collect, health, healthcare, transmit, hardware, wireless_sensor, patient, signal, measurement
Topic 7- Deep Learning	Image, accuracy, prediction, detection, deep_learning, video, classification, feature machine_learning, camera, detect, recognition, congestion, neural_network
Topic 8- Smart Manufacturing	Machine, Production, embed, robot, consumer, world, factory, standard, operational, communicate, digital, evolve, smart_factory, enterprise, home, vision
Topic 9- Computation Offloading	Task, offload, computation_offload, mobile, edge_server, load_balance, task_offloading, assign, battery, priority, offloading_decision, computation_intensive_reward, profit

Table 4 continued:

Topic 10-Robotics	Project, sensing, robotic, centric, data_acquisition, charge, augmented, communication_protocol, drone, plug, logic, chip, high_reliability

Results of the lemmatized model are similar in identifying most of the topics semantically can be related. For example, the wearable device trend identified in the non-lemmatized model is an important technology for real-time monitoring tasks. Case studies using data from wearable technologies are also identified in the SLR in Chapter 4. Also non-lemmatized model could not have identified a topic related to robotics technologies as Topic 10. Also, we identified case studies conducted using robotics data in the SLR, and we know it is an important technology trend. Therefore, the author decided to utilize both models for findings.

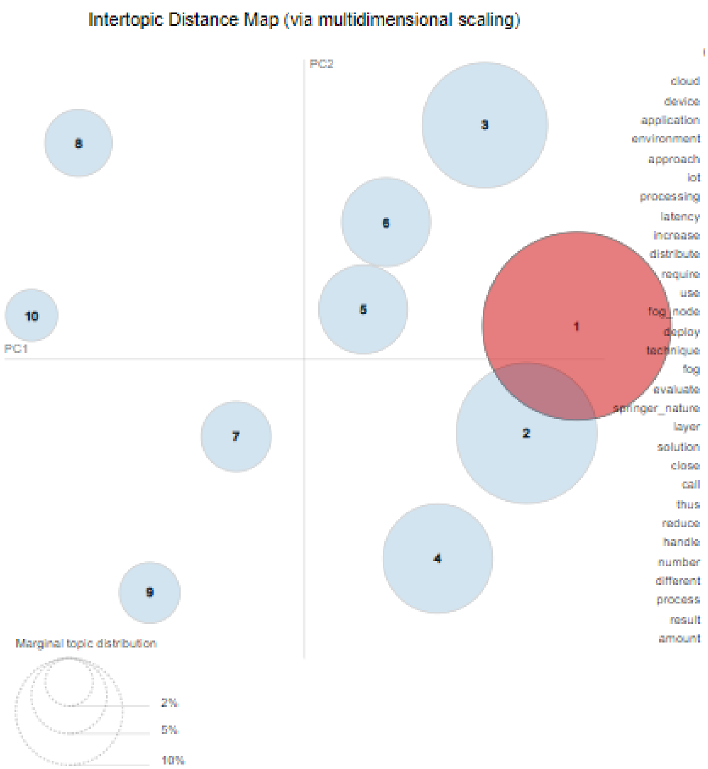

Figure 9- Intertopic Distance Map of Lemmatized LDA

5.5 BERTopic (Attention is All Researchers Need)

Transformer-based deep learning models provided incredible advances in NLP tasks. They adopt specifically the attention mechanism from Recurrent Neural Networks (RNN) (Vaswani et al., 2017), enabling the model to handle long sequences of words without overfitting to the last appearing ones. In other words, they can process multiple words, therefore a complete sentence at the same time. Pre-trained in large textual databases; GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers) provided observable advances in NLP tasks. BERTopic is a topic modeling algorithm with several steps explained below that utilizes sentence-BERT for embedding, UMAP (Uniform Manifold Approximation and Projection for Dimension Reduction) for dimensionality reduction, HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) and c-TF-IDF (Term frequency-inverse document frequency) for clustering (Grootendorst, 2022). BERTopic is a more computationally complex algorithm when compared to LDA, therefore requires more time and computational resources to get results. It has a relatively easy implementation and supports guided, hierarchical, online, and dynamic topic modeling, by leveraging, class-based TF-IDF.

5.5.1 Algorithm and components

First, documents are transformed to vectors using sentence-transformers, a pre-trained using sentence BERT model “all-MiniLM-L6-v2”. Sentence BERT is a modified version of BERT to do semantic similarity clustering tasks. Modifications include Siamese and triplet network structures to compare similarities of clusters, otherwise would require pairwise comparison among all sentences, which is inconvenient computationally (Reimers and Gurevych, 2019). Another approach for this step; Top2vec representation for converting topics to vectors by word semantic embedding and also does not require discarding stop-words or lemmatization and automatically finds the number of topics (Angelov, 2020). BERTopic implements sentence BERT and compares it to the original paper (Grootendorst, 2022).

Following the embeddings, dimensionality reduction has been conducted as (Allaoui et al., 2020) state dimensionality reduction improves clustering performance. UMAP technique is selected, as it has been shown to outperform popular methods such as t-SNE and PCA (McInnes et al., 2020). The novel technique, based on Riemannian geometry and algebraic topology is discussed as superior to t-SNE in terms of preserving the global structure of the data. While (Kobak and Linderman, 2021) argue there is no evidence UMAP has proven advantage therefore PCA and t-SNE can also be used in BERTopic for dimensionality reduction.

After reducing the dimensions, the density-based clustering method HDBSCAN is used. For assessing the clustering, TF-IDF scores are used as a measure of the relevance of a word in a cluster. TF-IDF combines term frequency and inverse document frequency. Inverse document frequency indicates the amount of information a term provides to a document. Initially, all documents are assigned in a cluster, then iteratively calculating TF-IDF by assigning documents to clusters by the importance of words for each topic and merging, until the specified number of topics.

Dynamic topic modeling introduced by (Blei and Lafferty, 2006) enables analyzing changes in documents over time. This feature is modeled using class-based TF-IDF in topics. To achieve the dynamic nature, BERTopic multiples timestamp with term frequency of documents. This enables dynamic topic modeling and analyzing how trends change over time.

TF-IDF can be formulated as:

$$W(t,d) = TF(t,d) \cdot \log\frac{N}{df_t} \quad (1)$$

Where the variables are:

- N is the number of documents,
- D is an index for each document
- T is an index for each term
- tf $_{(t,d)}$ is a function that returns the term frequency, the number of times term t appears in document d,
- df $_{(t)}$ is the document frequency of term t.
- Inverse document frequency (IDF) is the logarithmic part of the statement, the logarithm of the number of documents N divided by document frequency df_t.

5.5.2 BERTopic Topic Model

Because BERTopic utilizes pre-trained BERT which is trained using real text, lemmatization is not used. Literature indicates there is no preprocessing needed for BERTopic (Kukushkin et al., 2022). In Figure 10 below, the number of topics is detected automatically by BERTopic. For achieving a better model, BERTopic allows merging topics, this is done manually by the researcher, investigating the similarity scores of keywords and the intertopic distance map. The similarity in topics is based on cTF-IDF scores.

Figure 10- Two topics to be merged in BERTopic Intertopic Distance Map

For example, two topics after the initial run, topics 5 and 11 with common keywords such as in this case: “security”, “attack”, “detection”, and the other “privacy”, “scheme”, and “encryption” are merged into one topic. This way, each topic is investigated after training and merged with a similar topic when necessary. By default, the BERTopic algorithm detected 50 topics capturing technology trends more in detail than LDA. By merging, these topics are reduced to 30. Experiments have been made with several minimum topic sizes and the number of topic values. Since LDA already gave highlights of edge computing trends, BERTopic enabled more detailed technological trends and relations to application areas.

Table 5: Topics and keywords of the BERTopic Topic Model

Topics	**Keywords**
T1- Internet of Vehicles	Vehicles (vehicular, vehicles), UAV, offloading, computing
T2- Blockchain-Security	Blockchain, security, data, privacy, scheme, attacks
T3- Manufacturing	Manufacturing, production, data, system, real, time
T4- Smart Grid	Grid, power, energy, smart, data, load, system, electricity
T5- Healthcare	Healthcare, health, medical, data, patient, monitoring
T6- Deep Learning	Deep, learning, neural, AI, devices, inference, dnn, models
T7- Blockchain-Management	Blokchain, miners, auction, mining, mobile, mec, resource, game

Table 5 Continued:

T8- Federated Learning	Federated, learning, privacy, data, local, distributed
T9- IoV-Security	Vehicles, blockchain, security, IoV, scheme, privacy, authentication
T10- Video Surveillance	Vide, surveillance, detection, object, camera, frame, accuracy, real
T11- Trust	Trust, attacks, devices, ddos, security, trustworthiness, malicious
T12- MEC	MEC, mobile, security, network, access, services, key, service
T13- Smart City	Smart, city, urban, technologies, data, citizens, big, service
T14- Agriculture	Agriculture, crop, farmers, farming, data, soil, plant
T15- Routing	Network, routing, energy, cluster, sensor, NFV(network functions virtualization), lifetime, VNF (virtual network function)
T16- Edge Databases (storage)	Data, processing, database, storage, distributed, prediction, real-time
T17- Robotics	Robotics, robots, slam (Simultaneous localization and mapping), cloud, localization, system, computing
T18- Caching	Caching, cache, query, network, hit, user, strategy, replacement, algorithm, popularity
T19- Software Defined Networking	Network, slicing, SDN, NFV, access, generation, radio, technologies, spectrum,
T20- Augmented Reality	AR, Video, VR, reality, augmented, streaming, caching, MAR (Mobile augmented reality), content

Table 5 Continued:

T21- Facial Emotion Recognition	Recognition, emotion, human, face, activity, facial, accuracy, system
T22- Smart Grid	Grid, smart, aggregation, privacy, scheme, data, security, meters, preserving, electricity
T23- Anomaly Detection	Anomaly, detection, data, sensor, time, detect, abnormal, source, real
T24- Air Quality Monitoring	Air, quality, pollution, monitoring, indoor, system time, sensors
T25- Water Drought Monitoring	Water, drought, monitoring, severity, quality, flood, using, system, prediction
T26- Traffic Detection	Traffic, detection, vehicle, road, object, lane, time, real, prediction, accidents
T27- Smart Car Parking	Parking, smart, system, car, slot, finding, space, vehicle, city
T28- Digital Twin	Digital, twin, manufacturing, physical, virtual, machine, construction, data, wireless
T29- Kubernetes	Kubernetes, container, scheduler, pod, orchestration, autoscaling, cluster, resource

Notice “real” and “time” words occurring in the same topics, apparently indicating real-time. Topics 3 and 4, namely smart grid and manufacturing also have common tokens and are intertwined in the intertopic distance map similarly in Figure 8. However, these topics reflect two different big market trends so they are not merged into the same topic. Below, c-TF-IDF scores are given for terms and topics in Figure 11. A higher score on the word means a higher probability of being a part of that topic.

Figure 11- BERTopic Topic Word Scores

In the results, blockchain technology is used for two different business objectives, security applications, and resource management applications, identified as topics 2 and 7. There is a bilateral relationship between the convergence of edge computing and blockchain technologies. One main challenge of blockchain technology is the scalability of computing resources, while a big challenge of edge architectures is managing decentralized devices securely. Keywords of topic 2, “security”, “data”, “privacy”, and “ attacks” make it obvious on using blockchain for increasing security. Keywords identified in topic seven: “miners”, “auction”, “mobile”,” resource”, and “game” refer to game theoretic systems for managing decentralized computing resources. Although this thesis does not investigate how edge technologies can be used in blockchain applications such as managing computing resources for mining cryptocurrencies, topic 7, blockchain management is taken as a technological trend for managing edge nodes, similar to SDN and NFV.

5.6 Dynamic Topic Modeling Results

In recent literature, several papers discussed BERTopic gives more human-interpretable results than LDA (Egger and Yu, 2022; Kukushkin et al., 2022). This is also backed by our results for our problem and dataset.

Dynamic topic modeling enables one to see how edge applications and market trends changed over time. Although it is inconvenient to analyze all the topics by time in a single plot, the dynamic plot in the BERTopic library enables one to select multiple topics to analyze them. See the topic numbers and keywords can be chosen from the right-hand side. Below, Topics 16, 17, 18, 21, 28, and 29 are shown with the relative trend to each other. For example, it can be seen that a sharp increase in blockchain-related research started to decline from 2021 to 2022, while cloudlet and Kubernetes technological trends continued to increase.

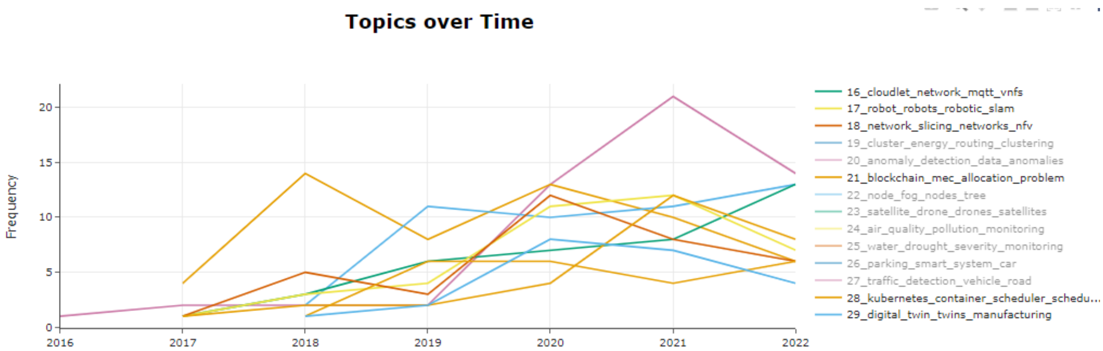

Figure 12- Dynamic Topics Trend Analysis

Each topic identified from the dynamic topic model is investigated and as explained in Chapter 5.8, time series information is used for creating a technology roadmap.

5.7 Cross-chapter Discussion

Recall in the discussion chapter 4.3, the following trends were identified as results of the SLR:

- Latency is a common concern among edge applications in IIoT.
- Depending on the complexity of the computation business case requires, e.g., a deep learning task or real-time monitoring, literature presents offloading mechanisms or tailored approaches for efficiently using resources.
- Open-source tools have been commonly used to develop edge architectures.
- Blockchain has been incorporated into IIoT networks for security.
- Approaches such as SDN, containerization, and computational offloading algorithms are used for resource utilization and energy efficiency.

Findings in the topic modeling section validate these trends by identifying blockchain and security-related topics, both in LDA and BERTopic, deep learning applications trends, SDN and containerization keywords for computational offloading topics, energy efficiency topics, and Kubernetes topic is identified in BERTopic as topic 29.

5.8 Data-Integrated Technology Roadmap

Findings of the topic modeling are integrated to form an edge computing roadmap. Links and relationships between the roadmap components are formed both using terms occurring in the same topics, also from the manually investigated papers from the SLR and expert opinions. As explained in Section 2, using edge computing is a data-related issue therefore it should be planned and analyzed with a data-centric approach (Kayabay et al., 2020). Therefore, a data-integrated technology roadmap for edge computing is proposed. Data sources and processes are linked with market and technology layers. Some linkages and background information are explained in the technology roadmap.

5.8.1 Smart Manufacturing and Real-Time Monitoring

Smart manufacturing use cases and data sources are briefly explained in Chapter 2. In addition to those, technology trends identified by the topic model that are related to smart manufacturing are robotics, digital twins, anomaly detection, and deep learning applications applied to manufacturing data. Notably, LDA results include ‘detection’, ‘recognition’, ‘classification’, ‘accuracy’ tokens in the smart manufacturing topic.

Although it is included in the monitoring-related topic in the LDA results, “wearable technologies” is another technology trend also used in smart manufacturing. Regarding how wearables enable real-time monitoring and smart manufacturing market trends, for example, monitoring the amounts of harmful gas spread, radiation, or dangerous lightning during operations in real-time can prevent work accidents. These can also be done using detectors on tools, trucks, or conveyors carrying semi materials. In contrast to monitoring and doing predictive maintenance analytics with data collected from machinery, performance, efficiency, and health status can also be used for workers using wearables. A case study conducted in Australia (Forkan et al., 2019) monitors and analyzes worker performance using data streams collected from wearable sensors and transmitted to edge computers of Raspberrypis using Bluetooth. Then captured data is forwarded to central cloud servers for machine learning, detecting the activities, and calculating KPIs.

Other examples are pumps with limited computing capabilities that can shut down when dangerous temperatures are exceeded, protecting equipment and personnel (Peng et al., 2019). In production lines of subassemblies, keeping track of KPIs as overall equipment effectiveness (OEE) is traditionally done by sampling and observations of production engineers. These KPIs can be provided in real-near time to required personnel in edge at the plant level, avoiding costs of working with low performance before detecting inefficiencies and decisions from the center are made. Similar examples can be given in monitoring supply chains or predictive maintenance.

5.8.2 Internet of Vehicles

Utilizing driving data enables applications such as predicting driver wakefulness and alarming for rests when needed. Real-time geographical information on vehicles reduces traffic congestion in the cities. Topic modeling results indicated smart car parking as a technological trend in this domain. Also, video surveillance is detected as an enabler of autonomous driving. Federated learning can be integrated for securing users' privacy while enabling the development of collaborative deep-learning models for autonomous driving (Lim et al., 2021). Autonomous vehicles can both use and produce data for geographical position, physical data such as velocity, tire pressure, parameters of engine data, accidents, and safety of the road (Xu et al., 2018). This way it can also be an enabler of smart logistics with less randomness in the supply chains. Where IoV and smart manufacturing meet; fleets of automated guided vehicles (AGV) or similar coordinated robots used in factory floors for packaging, palletizing, placing, and inspections having sensors with edge computing capabilities may not require human orchestration (Industrial Internet Consortium et al., 2016).

5.8.3 Smart Grid

Topic model results show the smart meter as an important enabler technology identified in the topic model. Topic models identify smart grid is a very close trend to smart manufacturing in terms of similarity scores of keywords. As stated in Chapter 4.3.1 real-time electricity consumption data from household and industrial plants, and production data from power plants are sources of data in a smart grid (Feng et al., 2021). Customers can monitor their energy consumption in real-time while maintaining security because end-user data are stored at the edge nodes, not in the cloud server (Feng et al., 2021). Also, see in the smart grid topic identified in the LDA results, shows the privacy of user electricity consumption is another challenge edge computing literature has been working on. Customers can monitor their energy consumption in real-time while maintaining security because end-user data are stored at the edge nodes, not in the cloud server (Feng et al., 2021). Balanced energy load in a grid can be achieved using these architectures in the smart grid with low-latency. (Okay and Ozdemir, 2016).

Edge computing in smart grids can make sustainable energy more dependable and useful. As wind speed and direction change in a wind farm, edge devices analyze data in real-time, optimize accordingly, and send preprocessed data to the central cloud, reducing communication time and data transfer costs. These data and optimization would help an energy company trade better terms in the market while avoiding wasting resources and increasing dependability on renewable energy, which is the most critical aspect reason still use fossil fuels.

5.8.4 Healthcare 4.0

Digital transformation of healthcare is referred to as Healthcare 4.0, smart healthcare, or digital healthcare in different domains. Healthcare statistics, patients diagnosed with a disease, analysis, and records from patients are examples of data leveraged by researchers, and pharmaceutical companies in the healthcare domain. Topic modeling results also show in the healthcare literature, privacy challenges of the patient data are addressed using integrating blockchain to these data, also processing using federated learning. Tomography images, or magnetic resonance imaging (MRI) data are processed with deep learning and tumors can be detected (Rieke et al., 2020).

5.8.5 Smart Farming

As explained in Chapter 2 with a case study (Sittón-Candanedo et al., 2019), the digital transformation of agriculture helps farmers work more efficiently, reliably, and sustainably. Data can be collected from RFIDs, sensors, barcodes for temperature, humidity, levels of CO_2 or other chemicals, and lightning for agriculture (O’Grady et al., 2019). Besides production, digital transformation in agriculture also enables transparency and traceability of the food across the supply chain for consumers. The main reason for monitoring food is to make instant decisions for possible flaws in products because they are quickly perishable. Data regarding packaging, transporting pallets, containers and can be traced through the supply chain, identifying products at different points. This way, edge computing provides a latency-sensitive ubiquitous product lifecycle monitoring environment in the food industry.

5.8.6 Environmental Monitoring

The covid-19 pandemic increased the trend of air quality monitoring. In Chapter 4, a university campus's air quality monitoring system (Kristiani et al., 2021). Data was collected from Arduino sensors across the campus containing temperature, humidity, PM10 (Particulate Matter), and PM5, regarding particles less than 10 and 5 µm in diameter. Another topic identified in BERTopic includes data processes water quality monitoring and water draught prediction. Dissolved oxygen is a key parameter for using water in agriculture, which can be monitored using edge architectures (Kuang et al., 2020). Draught prediction utilizes data such as water supply resources, meteorological conditions, soil moisture, drought-causing factors data from government agencies, and periodic rainfalls (Kaur and Sood, 2020).

5.8.7 Smart City

Smart city can be seen as a higher trend that healthcare, smart grid, agriculture, environmental monitoring, and smart transportation systems (IoV) enables. Lemmatized LDA topic model also presents "smart_grid" and "smart_city" bigram tokens in topic 3, so we infer the linkage between these trends. Also, BERTopic identified smart car parking trend as topic 27, which includes "city" as a token, which identifies a technology market trend mapping, added in the technology roadmap. All research and technologies for preventing privacy such as blockchain and federated learning are directly related to smart cities. As identified in all of the topic models and mentioned under 5.8.2, real-time traffic monitoring and analytics can be used in new data applications in smart cities, finding better routes, arranging timings of traffic lights, and forwarding vehicle loads of specific areas (Khan et al., 2020).

Other future applications in smart cities can be, creating security applications using video surveillance technology using data from security cameras (Markavathi and Kesavaraja, 2021, p. 8). Although some aspects of this use case can be seen as a violation of human freedom depending on how the automatic data is utilized and stored. Another example of data application is smart buildings or houses as parts of smart cities can monitor and control temperatures, gas levels, or humidity as discussed similarly in Chapter 5.8.6 (Abbas et al., 2018).

5.8.8 *Final Roadmap*

In the roadmap below shown in Figure 13, the first appearances of each trend item in the dynamic topic model are mapped as the year. In a column of a year, market trends, data applications, data sources, and enabling technologies are mapped. However, some technologies started being used after the start of the market trend. For example, edge computing applications in healthcare studies were present from 2014 so we see that trend under 2014. Federated learning technology trend in edge computing research, started being applied to healthcare data after 2019, therefore arrows are used for mapping those roadmap items. An arrow from technologies to data sources indicates the relationship identified from topic modeling, and that technology is frequently applied to that data source, hence in the data application as well, enabling the market trend in the same column. Similarly, an arrow from the data layer to market layer also indicates high similarity in the topic model and semantic causality relationship.

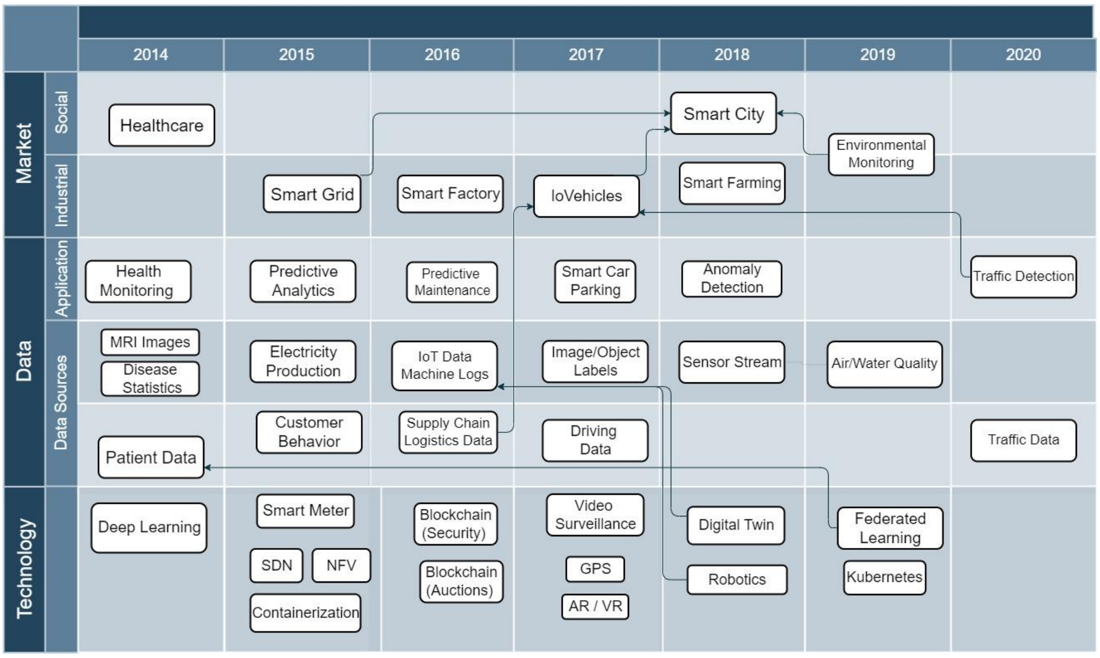

Figure 13- Edge Computing in IIoT Final Roadmap

From the mapping explanations in Chapter 5.8, technological trends of smart meters, blockchain security, video surveillance, digital twin, robotics, and federated learning are mapped to their market trends according to the topic model results and the most widely adopted application domains in the SLR. For example, it is obvious that smart meter technology is a key enable for smart grids, or digital twins are more likely to be used in smart manufacturing. However, other identified technology trends namely, deep learning, SDN, NFV, containerization, blockchain management for computational power, and Kubernetes, can be used in almost every market trend application, since they focus on edge architecture resource utilization improvements. Technologies such as SDN, containerization, and Kubernetes are use case dependent edge computing technology trends on their own and can be related to all edge architectures independent of the application domain and market trend. Examples of applications of those technologies in edge architectures are given in Chapter 4.2.1 and Chapter 4.2.3, and grouped regarding their business objectives such as increasing resource utilization and energy efficiency applications are discussed under Chapter 4. These examples are mainly in smart manufacturing and monitoring tasks considering the focus of SLR in Chapter 4. Therefore, these technology trends should not be seen as mapped only to the market trends under which they are located in the same column they appear.

5.9 Discussion: Additional Implications for Organizations

As stated in Chapter 6.3, this study proposes a high-level sectoral technology roadmap for edge computing and discusses the need for future research, especially at the organizational level. This high-sectoral roadmap brings us a limitation also stated in Chapter 6.3, the generality of the roadmap makes it less strategic (Kanama, 2013). Just as this study focuses on edge computing, data-driven technology roadmapping studies generally focus on science and technology related research. This is not a limitation due to the properties of the process, but the availability of large textual data (Kostoff and Schaller, 2001). One limitation of data-driven approaches is that they are subject to plentiful uncertainties regarding organizational matters and are rarely applied and validated in real-world settings (Park et al., 2020). The potential strategic benefits of the technology roadmapping process for an organization can be exploited by customizing it to company-specific internal and external factors (Lee and Park, 2005). Quantitative studies and our technology roadmap can be used as input for qualitative methods to conduct further technology forecasting as in hybrid methodology studies explained in Chapter 2.9.2. These studies combined text mining and similar quantitative approaches with qualitative approaches such as workshops (Noh et al., 2021), (Wang et al., 2018) or interviews with experts (Ozcan et al., 2022), (Ma et al., 2021), (Li et al., 2019).

Using the guidelines by (Phaal et al., 2004a) on customizing a fast-start T-Plan process, baseline implications regarding creating an organizational edge computing strategy are proposed, which can be detailed and validated with a case study in an organization in future work. Practitioners can discuss these topics consecutively in their workshops for an edge adoption and should customize workshops based on organizational structure. Below, recommended steps for edge computing practitioners and researchers from the socio-technical perspective on the IIoT environment are discussed and shown in Figure 14.

In the recently proposed data science roadmapping framework (Kayabay et al., 2022), the customized roadmapping process involves four main workshops, where strategical, data-related, technological, and organizational aspects are discussed, respectively. Similarly, this study proposes the steps below for fast starting a strategic edge computing planning.

1. Learn industry and market trends, presented in this study in Chapter 5, and discuss how a company can align. Analyze the current production workflows to find opportunities and set up a strategy for filling those gaps. Looking for examples of an edge architecture and forming a data-driven business process that utilizes the architecture may also be a more specific aim for the company for starting digitally transforming current business processes.

2. Discuss which business processes can be leveraged using edge, and how to exploit the value of real-time data analytics using an edge architecture. A company may want to achieve one of the following scenarios: real-time monitoring of several processes and KPIs, making forecasts with data from distributed network devices, detecting anomalies, tracing machinery for predictive maintenance, and sustaining these in the business processes.

3. The sufficiency of currently used technologies should be evaluated. Decide on whether to find a provider for a service level agreement or create the whole architecture using open-source tools. Resources like budget, time, human resources, relationship with academia, OT/IT domain knowledge can be also determined. Idle OT devices or computing resources may be used for prototypes.

4. Create necessary teams, determining suitable personnel according to the required skillset for each step depending on the type and size of the organization. Best practices, issues, and solutions in the industry/market should be investigated, especially security concerns edge computing brings. Top management support for the process and cooperation between business functions are essential. Representatives of partner companies or consultants may attend to this process to share know-how and speed up processes.

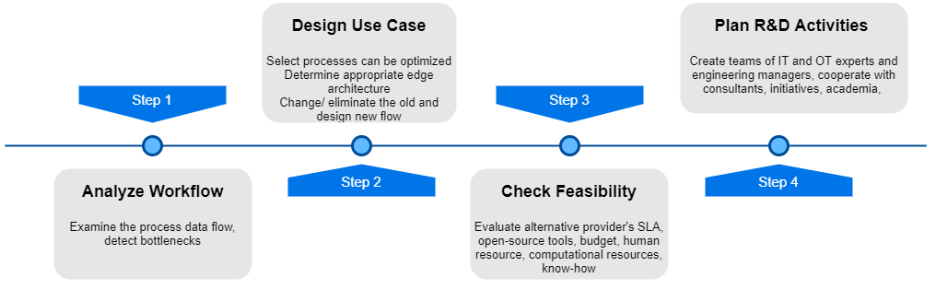

Figure 14- Proposed steps for adopting edge architectures

After identifying possible architectures, a team with OT and IT skills has to decide on the components and technological infrastructure. Consultants or partner company representatives can contribute depending on the experience of the personnel and the complexity of the architecture company decides to achieve. Working together with academia and companies with expertise is vital for understanding the best practices.

CHAPTER 6

CONCLUSION

6.1 Summary

A technology roadmap can be used for analyzing or forecasting technology trends, customer needs, relationships between alternative technologies, competitors, changes in market conditions, and developing strategies to meet the industry's technology needs (Lee and Park, 2005; Lu and Weng, 2018). This study investigates edge computing research and applications in IIoT and related environments under a sectoral-industrial scope and proposes a retrospective data-integrated edge computing technology roadmap. The retrospective roadmap served as a technological evaluation and social assessment tool. Also, it can be used as a starting point for other organizational or prospective planning studies by integrating to complement the data-integrated roadmapping activities (Kayabay et al., 2022).

Research on edge computing is investigated and a knowledge gap is identified there are no strategic investigations, but only demo applications as stated in detail in Chapter 1.2 problem statement and Chapter 2.10 literature summary and knowledge gaps. First, a manual SLR is conducted including papers with applications in a real-life IIoT setting. Technologies in edge architectures, data sources used, business objectives, and the application domain are described.

To create the roadmap, market and technology trends of using edge computing in IIoT environments are identified using topic modeling. LDA and BERTopic algorithms are applied to a dataset of academic publications. The results of topic models are then combined with the results of SLR and shown in the technology roadmap. Dynamic topic modeling approach enabled us to see how these trends started and changed over time, as shown in the horizontal axis of the technology roadmap. Generally, market and technological trends are identified from the topic model and were supporting findings of the SLR. Some market and technology trends identified in the SLR are then validated with the results of the topic models and discussed under the cross-chapter discussion in Chapter 5.7. Data sources and applications, on the other hand, were more use case dependent than market and technology trends, therefore could not be identified in topic models, generalized using examples from the SLR, and integrated into the roadmap manually by the researchers. The creation of the final roadmap is explained under Chapter 5.8 and presented with explanations of the data sources and applications used in that market trend, justifying market-data layer relationships in the roadmap.

6.2 Contributions

This thesis contributes to the edge computing literature by proposing the first sectoral edge computing technology roadmap. It reveals how the focus of edge computing research and technologies correlates with business objectives and application domains in IIoT.

It contributes to technology roadmapping literature by developing a new technology and social change assessment tool, proposing a data-integrated roadmap for sectoral planning of emerging technology, by integrating the data layer into the technology roadmap.

Methodologically, it contributes to exploratory technology forecasting and assessment literature by combining SLR and topic modeling into technology roadmapping. It evaluates and presents BERTopic and LDA algorithms for creating topic models. It compares lemmatized and non-lemmatized LDA models and discusses the results and effects of lemmatization in topic models under Chapter 5.4. It integrates dynamic topic modeling into technology roadmapping stated as future work in related studies (Feng et al., 2022; Kim and Geum, 2021). Dynamic topic model using BERTopic is explained in Chapter 5.5 and Chapter 5.6 explained how the time dimension from the results is integrated into the TRM. Chapter 5.9 discusses implications for organizations on how to make a more strategic and customized edge computing planning by tailoring a workshop-based roadmapping process to their requirements.

Therefore, this study used data-driven technology roadmapping to bridge the sociotechnical knowledge gaps in the edge computing domain.

6.3 Limitations and Future Work

Chapter 5.9 stated technology roadmaps can serve different scopes of planning such as organizational, industrial, or national. However as the purpose of the roadmap becomes more general, the roadmap is less strategic (Kanama, 2013), therefore future research can use roadmapping to be incorporated at the organizational level.

Quite often, technology forecasting is incorporated into a roadmapping activity (Lee and Park, 2005), which might indicate it is a tool for technology forecasting, however as can also be seen in other data-driven roadmapping studies presented in Chapter 2.9.2, it provides a mechanism to help experts make the forecasting (Garcia and Bray, 1997). In this study, the technology roadmap assesses the current situation of edge computing in IIoT, without making any forecasts for the future. Also, further qualitative studies can be integrated such as workshops or expert interviews can be integrated into the findings similar to some studies explained in Chapter 2.9.2 for making the roadmap prospective rather than retrospective, therefore the roadmap can be used as a technology forecasting tool as well. Chapter 5.9 discusses how a potential workshop-based approach could be customized for edge architectures.

Future work can conduct a case study for developing an edge computing adoption strategy for industrial organizations. Also, data-integrated roadmapping studies (Kayabay et al., 2022) and (Han and Geum, 2020) can be used for more data-centric strategy development at the organizational level, since the motivations of using edge computing are shaped around the characteristics of data available.

In the final edge computing technology roadmap in Chapter 5.8, columns reflect the market, data, and technology trends that are correlated to each other, however, most of the technology trends identified can be used in all of the market trends, so are not individually stated as relationships. For example, it is obvious that smart meter technology enables smart grids or digital twins to be more likely used in smart manufacturing. However, technologies such as SDN, containerization, and Kubernetes are edge computing technology trends on their own and can be related to all edge architectures independent of the application domain and market trend.

As for data sources of technological and market trends, patents can be utilized as the data source with well-established database search results in patent databases. For market trends, sources from grey literature such as commercial edge solutions, success stories, or market research company reports, can be used. Also, other quantitative approaches can be used for finding further and more detailed links between market and technology trends.

REFERENCES

Abbas, N., Zhang, Y., Taherkordi, A., Skeie, T., 2018. Mobile Edge Computing: A Survey. IEEE Internet of Things Journal 5, 450–465. https://doi.org/10.1109/JIOT.2017.2750180

Accenture, 2023. What Is Edge Computing & Why Is It Important? | Accenture [WWW Document]. URL https://www.accenture.com/us-en/insights/cloud/edge-computing-index (accessed 2.24.23).

Al-Ali, A.G., Phaal, R., 2019. Design Sprints for Roadmapping an Agile Digital Transformation, in: 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). Presented at the 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), pp. 1–6. https://doi.org/10.1109/ICE.2019.8792597

Al-Ali, A.G., Phaal, R., Sull, D., 2020. Deep Learning Framework for Measuring the Digital Strategy of Companies from Earnings Calls, in: Proceedings of the 28th International Conference on Computational Linguistics. Presented at the COLING 2020, International Committee on Computational Linguistics, Barcelona, Spain (Online), pp. 927–935. https://doi.org/10.18653/v1/2020.coling-main.80

Allaoui, M., Kherfi, M.L., Cheriet, A., 2020. Considerably Improving Clustering Algorithms Using UMAP Dimensionality Reduction Technique: A Comparative Study, in: El Moataz, A., Mammass, D., Mansouri, A., Nouboud, F. (Eds.), Image and Signal Processing, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 317–325. https://doi.org/10.1007/978-3-030-51935-3_34

AllTheResearch, 2021. Industrial IoT (IIoT) Market by Component (Solution, Services, Platform), End-Use Vertical | COVID-19 Impact Analysis | Top Trends and Forecast 2021 - AllTheResearch [WWW Document]. URL https://www.alltheresearch.com/report/226/iiot-market (accessed 12.20.21).

Angelov, D., 2020. Top2Vec: Distributed Representations of Topics. https://doi.org/10.48550/arXiv.2008.09470

Ashton, K., 2009. That “Internet of Things” Thing. RFID Journal LLC.

Asmussen, C.B., Møller, C., 2019. Smart literature review: a practical topic modelling approach to exploratory literature review. Journal of Big Data 6, 93. https://doi.org/10.1186/s40537-019-0255-7

Atieh, A.T., 2021. The Next Generation Cloud technologies: A Review On Distributed Cloud, Fog And Edge Computing and Their Opportunities and Challenges. ResearchBerg Review of Science and Technology 1, 1–15.

Atzeni, D., Bacciu, D., Mazzei, D., Prencipe, G., 2022. A Systematic Review of Wi-Fi and Machine Learning Integration with Topic Modeling Techniques. Sensors 22, 4925. https://doi.org/10.3390/s22134925

Balasubramanian, S., M.D., 2022. Philips Is Betting Big On Digital Health [WWW Document]. Forbes. URL https://www.forbes.com/sites/saibala/2022/03/22/philips-is-betting-big-on-digital-health/ (accessed 4.19.23).

Banijamali, A., Pakanen, O.-P., Kuvaja, P., Oivo, M., 2020. Software architectures of the convergence of cloud computing and the Internet of Things: A systematic literature review. Information and Software Technology 122, 106271. https://doi.org/10.1016/j.infsof.2020.106271

Barzegaran, M., Desai, N., Qian, J., Tange, K., Zarrin, B., Pop, P., Kuusela, J., 2020. Fogification of electric drives: An industrial use case, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Presented at the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 77–84. https://doi.org/10.1109/ETFA46521.2020.9212010

Bayar, A., Şener, U., Kayabay, K., Eren, P.E., 2023. Edge Computing Applications in Industrial IoT: A Literature Review, in: Bañares, J.Á., Altmann, J., Agmon Ben-Yehuda, O., Djemame, K., Stankovski, V., Tuffin, B. (Eds.), Economics of Grids, Clouds, Systems, and Services, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 124–131. https://doi.org/10.1007/978-3-031-29315-3_11

Blei, D.M., Lafferty, J.D., 2006. Dynamic topic models, in: Proceedings of the 23rd International Conference on Machine Learning, ICML '06. Association for Computing Machinery, New York, NY, USA, pp. 113–120. https://doi.org/10.1145/1143844.1143859

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Bose, S.K., Kar, B., Roy, M., Gopalakrishnan, P.K., Basu, A., 2019. ADEPOS: anomaly detection based power saving for predictive maintenance using edge computing, in: Proceedings of the 24th Asia and South Pacific Design Automation Conference. Presented at the ASPDAC '19: 24th Asia and South Pacific Design Automation Conference, ACM, Tokyo Japan, pp. 597–602. https://doi.org/10.1145/3287624.3287716

Carlos Brava, Henrik Backström, 2020. Whitepaper on Edge computing deployment strategies [WWW Document]. URL https://www.ericsson.com/en/reports-and-papers/white-papers/edge-computing-and-deployment-strategies-for-communication-service-providers (accessed 12.20.21).

Carvalho, G., Cabral, B., Pereira, V., Bernardino, J., 2021. Edge computing: current trends, research challenges and future directions. Computing 103, 993–1023. https://doi.org/10.1007/s00607-020-00896-5

Cetindamar, D., Phaal, R., 2017. Technology Management: Activities and Tools. Bloomsbury Publishing.

Chalapathi, G.S.S., Chamola, V., Vaish, A., Buyya, R., 2021. Industrial Internet of Things (IIoT) Applications of Edge and Fog Computing: A Review and Future Directions, in: Chang, W., Wu, J. (Eds.), Fog/Edge Computing For Security, Privacy, and Applications, Advances in Information Security. Springer International Publishing, Cham, pp. 293–325. https://doi.org/10.1007/978-3-030-57328-7_12

Chang, J., Gerrish, S., Wang, C., Boyd-graber, J., Blei, D., 2009. Reading Tea Leaves: How Humans Interpret Topic Models, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Chekired, D.A., Khoukhi, L., Mouftah, H.T., 2018. Industrial IoT Data Scheduling Based on Hierarchical Fog Computing: A Key for Enabling Smart Factory. IEEE Transactions on Industrial Informatics 14, 4590–4602. https://doi.org/10.1109/TII.2018.2843802

Chen, B., Wan, J., Celesti, A., Li, D., Abbas, H., Zhang, Q., 2018. Edge Computing in IoT-Based Manufacturing. IEEE Communications Magazine 56, 103–109. https://doi.org/10.1109/MCOM.2018.1701231

Chen, C.-H., Lin, M.-Y., Liu, C.-C., 2018. Edge Computing Gateway of the Industrial Internet of Things Using Multiple Collaborative Microcontrollers. IEEE Network 32, 24–32. https://doi.org/10.1109/MNET.2018.1700146

Chuang, J., Manning, C.D., Heer, J., 2012. Termite: visualization techniques for assessing textual topic models, in: Proceedings of the International Working Conference on Advanced Visual Interfaces. Presented at the AVI'12: International Working Conference on Advanced Visual Interfaces, ACM, Capri Island Italy, pp. 74–77. https://doi.org/10.1145/2254556.2254572

Dahmen-Lhuissier, S., n.d. ETSI Multi-access Edge Computing starts second phase and renews leadership team [WWW Document]. ETSI. URL https://www.etsi.org/newsroom/news/1180-2017-03-news-etsi-multi-access-edge-computing-starts-second-phase-and-renews-leadership-team (accessed 2.2.23).

Daim, T.U., Rueda, G., Martin, H., Gerdsri, P., 2006. Forecasting emerging technologies: Use of bibliometrics and patent analysis. Technological Forecasting and Social Change, Tech Mining: Exploiting Science and Technology Information Resources 73, 981–1012. https://doi.org/10.1016/j.techfore.2006.04.004

Denzler, P., Ruh, J., Kadar, M., Avasalcai, C., Kastner, W., 2020. Towards Consolidating Industrial Use Cases on a Common Fog Computing Platform, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Presented at the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 172–179. https://doi.org/10.1109/ETFA46521.2020.9211885

Desouza, K. (Ed.), 2005. New Frontiers of Knowledge Management. Palgrave Macmillan, Houndmills, Basingstoke ; New York.

Dilley, J., Maggs, B., Parikh, J., Prokop, H., Sitaraman, R., Weihl, B., 2002. Globally distributed content delivery. IEEE Internet Comput. 6, 50–58. https://doi.org/10.1109/MIC.2002.1036038

Dissel, M.C., Phaal, R., Farrukh, C.J., Probert, D.R., 2009. Value Roadmapping. Research-Technology Management 52, 45–53. https://doi.org/10.1080/08956308.2009.11657599

Djemame, K., 2021. Energy Efficiency in Edge Environments: a Serverless Computing Approach, in: Tserpes, K., Altmann, J., Bañares, J.Á., Agmon Ben-Yehuda, O., Djemame, K., Stankovski, V., Tuffin, B. (Eds.), Economics of Grids, Clouds, Systems, and Services, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 181–184. https://doi.org/10.1007/978-3-030-92916-9_15

Egger, R., Yu, J., 2022. A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. Frontiers in Sociology 7.

Fabio Giust, Gianluca Verin, 2018. MEC Deployments in 4G and Evolution Towards 5G. ETSI Whitepaper 24.

Feng, C., Wang, Y., Chen, Q., Ding, Y., Strbac, G., Kang, C., 2021. Smart grid encounters edge computing: opportunities and applications. Advances in Applied Energy 1, 100006. https://doi.org/10.1016/j.adapen.2020.100006

Feng, L., Zhao, W., Wang, J., Lin, K.-Y., Guo, Y., Zhang, L., 2022. Data-Driven Technology Roadmaps to Identify Potential Technology Opportunities for Hyperuricemia Drugs. Pharmaceuticals 15, 1357. https://doi.org/10.3390/ph15111357

Ferrari, P., Rinaldi, S., Sisinni, E., Colombo, F., Ghelfi, F., Maffei, D., Malara, M., 2019. Performance evaluation of full-cloud and edge-cloud architectures for Industrial IoT anomaly detection based on deep learning, in: 2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0 IoT). Presented at the 2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0 IoT), pp. 420–425. https://doi.org/10.1109/METROI4.2019.8792860

Forkan, A., Montori, F., Georgakopoulos, D., Jayaraman, P.P., Yavari, A., Morshed, A., 2019. An Industrial IoT Solution for Evaluating Workers’ Performance Via Activity Recognition. https://doi.org/10.1109/ICDCS.2019.00139

Fruhwirth, M., Ropposch, C., Pammer-Schindler, V., 2020. Supporting Data-Driven Business Model Innovations: A structured literature review on tools and methods. Journal of Business Models 8, 7–25. https://doi.org/10.5278/ojs.jbm.v8i1.3529

Fu, J.-S., Liu, Y., Chao, H.-C., Bhargava, B.K., Zhang, Z.-J., 2018. Secure Data Storage and Searching for Industrial IoT by Integrating Fog Computing and Cloud Computing. IEEE Transactions on Industrial Informatics 14, 4519–4528. https://doi.org/10.1109/TII.2018.2793350

Garcia, M.L., Bray, O.H., 1997. Fundamentals of technology roadmapping (No. SAND-97-0665). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States). https://doi.org/10.2172/471364

Gartner, 2018. What Edge Computing Means For Infrastructure And Operations Leaders [WWW Document]. Gartner. URL https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders (accessed 12.20.21).

Garza Ramos, A., Daim, T., Gaats, L., Hutmacher, D.W., Hackenberger, D., 2022. Technology roadmap for the development of a 3D cell culture workstation for a biomedical industry startup. Technological Forecasting and Social Change 174, 121213. https://doi.org/10.1016/j.techfore.2021.121213

Geum, Y., Lee, H., Lee, Y., Park, Y., 2015. Development of data-driven technology roadmap considering dependency: An ARM-based technology roadmapping. Technological Forecasting and Social Change 91, 264–279. https://doi.org/10.1016/j.techfore.2014.03.003

Ghazi Vakili, M., Demartini, C., Guerrera, M., Montrucchio, B., 2019. Open Source Fog Architecture for Industrial IoT Automation Based on Industrial Protocols, in: 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC). Presented at the 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), pp. 570–578. https://doi.org/10.1109/COMPSAC.2019.00088

Gokalp, M.O., Kayabay, K., Akyol, M.A., Eren, P.E., Koçyiğit, A., 2016. Big Data for Industry 4.0: A Conceptual Framework, in: 2016 International Conference on Computational Science and Computational Intelligence (CSCI). Presented at the 2016 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 431–434. https://doi.org/10.1109/CSCI.2016.0088

Gregory, M.J., 1995. Technology Management: A Process Approach. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 209, 347–356. https://doi.org/10.1243/PIME_PROC_1995_209_094_02

Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure.

Güven, E.Y., Çamurcu, A.Y., 2018. Edge Computing Security Application: Kılıç, in: 2018 3rd International Conference on Computer Science and Engineering (UBMK). Presented at the 2018 3rd International Conference on Computer Science and Engineering (UBMK), pp. 248–253. https://doi.org/10.1109/UBMK.2018.8566521

Hamm, A., Willner, A., Schieferdecker, I., 2019. Edge Computing: A Comprehensive Survey of Current Initiatives and a Roadmap for a Sustainable Edge Computing Development.

Han, M., Geum, Y., 2020. Roadmapping for Data: Concept and Typology of Data-Integrated Smart-Service Roadmaps. IEEE Transactions on Engineering Management 1–13. https://doi.org/10.1109/TEM.2020.3013295

Ho, J.-Y., Lee, K., 2022. Customization of Roadmap: Human-in-the-Loop Topic Modelling Approach, in: 2022 Portland International Conference on Management of Engineering and Technology (PICMET). Presented at the 2022 Portland International Conference on Management of Engineering and Technology (PICMET), pp. 1–13. https://doi.org/10.23919/PICMET53225.2022.9882893

Hoffman, M., Bach, F., Blei, D., 2010. Online Learning for Latent Dirichlet Allocation, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Hoyle, A., Goel, P., Hian-Cheong, A., Peskov, D., Boyd-Graber, J., Resnik, P., 2021. Is Automated Topic Model Evaluation Broken? The Incoherence of Coherence, in: Advances in Neural Information Processing Systems. Curran Associates, Inc., pp. 2018–2033.

Hu, Y.C., Sabella, D., Sprecher, N., Young, V., 2015. Mobile Edge Computing A key technology towards 5G.

IBM, 2020. What Is Edge Computing [WWW Document]. URL https://www.ibm.com/cloud/what-is-edge-computing (accessed 2.28.23).

Inaam ul Haq, M., Li, Q., Hou, J., 2022. Analyzing the Research Trends of IoT Using Topic Modeling. The Computer Journal 65, 2589–2609. https://doi.org/10.1093/comjnl/bxab091

Industrial Internet Consortium, Tseng Mitch, Todd Edmunds, Canaran, L., 2016. Introduction to Edge Computing in IIoT 19.

Iorga, M., Feldman, L., Barton, R., Martin, M.J., Goren, N., Mahmoudi, C., 2018. Fog computing conceptual model (No. NIST SP 500-325). National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.500-325

Jeong, Y., Jang, H., Yoon, B., 2021. Developing a risk-adaptive technology roadmap using a Bayesian network and topic modeling under deep uncertainty. Scientometrics 126, 3697–3722. https://doi.org/10.1007/s11192-021-03945-8

Jeong, Y., Yoon, B., 2015. Development of patent roadmap based on technology roadmap by analyzing patterns of patent development. Technovation, Opportunity Recognition and Creation 39–40, 37–52. https://doi.org/10.1016/j.technovation.2014.03.001

Jin, G., Jeong, Y., Yoon, B., 2015. Technology-driven roadmaps for identifying new product/market opportunities: Use of text mining and quality function deployment. Advanced Engineering Informatics 29, 126–138. https://doi.org/10.1016/j.aei.2014.11.001

Kanama, D., 2013. Development of Technology Foresight: Integration of Technology Roadmapping and the Delphi Method, in: Moehrle, M.G., Isenmann, R., Phaal, R. (Eds.), Technology Roadmapping for Strategy and Innovation. Springer, Berlin, Heidelberg, pp. 151–171. https://doi.org/10.1007/978-3-642-33923-3_10

Kaur, A., Sood, S.K., 2020. Artificial Intelligence-Based Model For Drought Prediction and Forecasting. The Computer Journal 63, 1704–1712. https://doi.org/10.1093/comjnl/bxz105

Kaur, K., Garg, S., Aujla, G.S., Kumar, N., Rodrigues, J.J.P.C., Guizani, M., 2018. Edge Computing in the Industrial Internet of Things Environment: Software-Defined-Networks-Based Edge-Cloud Interplay. IEEE Communications Magazine 56, 44–51. https://doi.org/10.1109/MCOM.2018.1700622

Kayabay, K., Gokalp, M.O., Gokalp, E., Eren, P.E., Kocyigit, A., 2020. Data Science Roadmapping: Towards an Architectural Framework, in: 2020 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD). Presented at the 2020 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), IEEE, Marrakech, Morocco, pp. 1–6. https://doi.org/10.1109/ICTMOD49425.2020.9380617

Kayabay, K., Gökalp, M.O., Gökalp, E., Erhan Eren, P., Koçyiğit, A., 2022. Data science roadmapping: An architectural framework for facilitating transformation towards a data-driven organization. Technological Forecasting and Social Change 174, 121264. https://doi.org/10.1016/j.techfore.2021.121264

Kerr, C., Phaal, R., 2020. Technology roadmapping: Industrial roots, forgotten history and unknown origins. Technological Forecasting and Social Change 155, 119967. https://doi.org/10.1016/j.techfore.2020.119967

Khan, L.U., Yaqoob, I., Tran, N.H., Kazmi, S.M.A., Dang, T.N., Hong, C.S., 2020. Edge-Computing-Enabled Smart Cities: A Comprehensive Survey. IEEE Internet of Things Journal 7, 10200–10232. https://doi.org/10.1109/JIOT.2020.2987070

Khan, W., Ahmed, E., Hakak, S., Yaqoob, I., Ahmed, A., 2019. Edge computing: A survey. Future Generation Computer Systems 97. https://doi.org/10.1016/j.future.2019.02.050

Kim, J., Geum, Y., 2021. How to develop data-driven technology roadmaps:The integration of topic modeling and link prediction. Technological Forecasting and Social Change 171, 120972. https://doi.org/10.1016/j.techfore.2021.120972

Kitchenham, B., 2004. Procedures for Performing Systematic Reviews 33.

Kitchenham, B., Charters, S., 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering 2.

Kobak, D., Linderman, G.C., 2021. Initialization is critical for preserving global data structure in both t-SNE and UMAP. Nat Biotechnol 39, 156–157. https://doi.org/10.1038/s41587-020-00809-z

Kostoff, R.N., Schaller, R.R., 2001. Science and technology roadmaps. IEEE Transactions on Engineering Management 48, 132–143. https://doi.org/10.1109/17.922473

Kristiani, E., Yang, C.-T., Huang, C.-Y., Wang, Y.-T., Ko, P.-C., 2021. The Implementation of a Cloud-Edge Computing Architecture Using OpenStack and Kubernetes for Air Quality Monitoring Application. Mobile Netw Appl 26, 1070–1092. https://doi.org/10.1007/s11036-020-01620-5

Kuang, L., Shi, P., Hua, C., Chen, B., Zhu, H., 2020. An Enhanced Extreme Learning Machine for Dissolved Oxygen Prediction in Wireless Sensor Networks. IEEE Access 8, 198730–198739. https://doi.org/10.1109/ACCESS.2020.3033455

Kuever, N., 2019. Cloud and edge computing in IoT: a short history. Bosch Digital Blog. URL https://blog.bosch-digital.com/cloud-and-edge-computing-for-iot-a-short-history/ (accessed 3.1.23).

Kukushkin, K., Ryabov, Y., Borovkov, A., 2022. Digital Twins: A Systematic Literature Review Based on Data Analysis and Topic Modeling. Data 7, 173. https://doi.org/10.3390/data7120173

Kumar, T., Harjula, E., Ejaz, M., Manzoor, A., Porambage, P., Ahmad, I., Liyanage, M., Braeken, A., Ylianttila, M., 2020. BlockEdge: Blockchain-Edge Framework for Industrial IoT Networks. IEEE Access 8, 154166–154185. https://doi.org/10.1109/ACCESS.2020.3017891

Kwon, K., Jun, S., Lee, Y.-J., Choi, S., Lee, C., 2022. Logistics Technology Forecasting Framework Using Patent Analysis for Technology Roadmap. Sustainability 14, 5430. https://doi.org/10.3390/su14095430

Lee, C.K.M., Huo, Y.Z., Zhang, S.Z., Ng, K.K.H., 2020. Design of a Smart Manufacturing System With the Application of Multi-Access Edge Computing and Blockchain Technology. IEEE Access 8, 28659–28667. https://doi.org/10.1109/ACCESS.2020.2972284

Lee, S., Park, Y., 2005. Customization of technology roadmaps according to roadmapping purposes: Overall process and detailed modules. Technological Forecasting and Social Change 72, 567–583. https://doi.org/10.1016/j.techfore.2004.11.006

Li, L., Ota, K., Dong, M., 2018. Deep Learning for Smart Industry: Efficient Manufacture Inspection System With Fog Computing. IEEE Transactions on Industrial Informatics 14, 4665–4673. https://doi.org/10.1109/TII.2018.2842821

Li, X., Li, D., Wan, J., Liu, C., Imran, M., 2018. Adaptive Transmission Optimization in SDN-Based Industrial Internet of Things With Edge Computing. IEEE Internet of Things Journal 5, 1351–1360. https://doi.org/10.1109/JIOT.2018.2797187

Li, X., Xie, Q., Daim, T., Huang, L., 2019. Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology. Technological Forecasting and Social Change 146, 432–449. https://doi.org/10.1016/j.techfore.2019.01.012

Li, X., Zhou, Y., Xue, L., Huang, L., 2015. Integrating bibliometrics and roadmapping methods: A case of dye-sensitized solar cell technology-based industry in China. Technological Forecasting and Social Change 97, 205–222. https://doi.org/10.1016/j.techfore.2014.05.007

Liang, F., Yu, W., Liu, X., Griffith, D., Golmie, N., 2020. Toward Edge-Based Deep Learning in Industrial Internet of Things. IEEE Internet of Things Journal 7, 4329–4341. https://doi.org/10.1109/JIOT.2019.2963635

Lim, W.Y.B., Huang, J., Xiong, Z., Kang, J., Niyato, D., Hua, X.-S., Leung, C., Miao, C., 2021. Towards Federated Learning in UAV-Enabled Internet of Vehicles: A Multi-Dimensional Contract-Matching Approach. IEEE Transactions on Intelligent Transportation Systems 22, 5140–5154. https://doi.org/10.1109/TITS.2021.3056341

Liu, F., Tang, G., Li, Y., Cai, Z., Zhang, X., Zhou, T., 2019. A Survey on Edge Computing Systems and Tools. Proc. IEEE 107, 1537–1562. https://doi.org/10.1109/JPROC.2019.2920341

Liu, Y., Lan, D., Pang, Z., Karlsson, M., Gong, S., 2021. Performance Evaluation of Containerization in Edge-Cloud Computing Stacks for Industrial Applications: A Client

Perspective. IEEE Open Journal of the Industrial Electronics Society 2, 153–168. https://doi.org/10.1109/OJIES.2021.3055901

Lu, H.-P., Weng, C.-I., 2018. Smart manufacturing technology, market maturity analysis and technology roadmap in the computer and electronic product manufacturing industry. Technological Forecasting and Social Change 133, 85–94. https://doi.org/10.1016/j.techfore.2018.03.005

Ma, T., Zhou, X., Liu, J., Lou, Z., Hua, Z., Wang, R., 2021. Combining topic modeling and SAO semantic analysis to identify technological opportunities of emerging technologies. Technological Forecasting and Social Change 173, 121159. https://doi.org/10.1016/j.techfore.2021.121159

Markavathi, J.N.P., Kesavaraja, D., 2021. Chapter 8 - Cloud/edge computing for smart cities, in: Krishnan, S., Balas, V.E., Julie, E.G., Robinson, Y.H., Kumar, R. (Eds.), Blockchain for Smart Cities. Elsevier, pp. 153–169. https://doi.org/10.1016/B978-0-12-824446-3.00011-9

MarketsandMarkets, 2021. Industrial IoT Market Size, Share and Trends Forecast to 2026 | MarketsandMarkets™ [WWW Document]. URL https://www.marketsandmarkets.com/Market-Reports/industrial-internet-of-things-market-129733727.html (accessed 12.20.21).

McInnes, L., Healy, J., Melville, J., 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. https://doi.org/10.48550/arXiv.1802.03426

Miao, H., Wang, Y., Li, X., Wu, F., 2022. Integrating Technology-Relationship-Technology Semantic Analysis and Technology Roadmapping Method: A Case of Elderly Smart Wear Technology. IEEE Transactions on Engineering Management 69, 262–278. https://doi.org/10.1109/TEM.2020.2970972

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013. Distributed Representations of Words and Phrases and their Compositionality. https://doi.org/10.48550/arXiv.1310.4546

National Research Council, 1987. Management of Technology: The Hidden Competitive Advantage. National Academies Press, Washington, D.C. https://doi.org/10.17226/18890

Nazarenko, A., Vishnevskiy, K., Meissner, D., Daim, T., 2022. Applying digital technologies in technology roadmapping to overcome individual biased assessments. Technovation 110, 102364. https://doi.org/10.1016/j.technovation.2021.102364

Noble, B.D., Satyanarayanan, M., Narayanan, D., Tilton, J.E., Flinn, J., Walker, K.R., 1997. Agile application-aware adaptation for mobility, in: Proceedings of the Sixteenth ACM Symposium on Operating Systems Principles, SOSP ’97. Association for

Computing Machinery, New York, NY, USA, pp. 276–287. https://doi.org/10.1145/268998.266708

Noh, H., Kim, K., Song, Y.-K., Lee, S., 2021. Opportunity-driven technology roadmapping: The case of 5G mobile services. Technological Forecasting and Social Change 163, 120452. https://doi.org/10.1016/j.techfore.2020.120452

O'Grady, M.J., Langton, D., O'Hare, G.M.P., 2019. Edge computing: A tractable model for smart agriculture? Artificial Intelligence in Agriculture 3, 42–51. https://doi.org/10.1016/j.aiia.2019.12.001

Okay, F.Y., Ozdemir, S., 2016. A fog computing based smart grid model, in: 2016 International Symposium on Networks, Computers and Communications (ISNCC). Presented at the 2016 International Symposium on Networks, Computers and Communications (ISNCC), pp. 1–6. https://doi.org/10.1109/ISNCC.2016.7746062

Oracle, 2023. What is the Internet of Things (IoT)? [WWW Document]. URL https://www.oracle.com/tr/internet-of-things/what-is-iot/ (accessed 12.20.21).

Oyekanlu, E., 2017. Predictive edge computing for time series of industrial IoT and large scale critical infrastructure based on open-source software analytic of big data, in: 2017 IEEE International Conference on Big Data (Big Data). Presented at the 2017 IEEE International Conference on Big Data (Big Data), pp. 1663–1669. https://doi.org/10.1109/BigData.2017.8258103

Ozcan, S., Homayounfard, A., Simms, C., Wasim, J., 2022. Technology Roadmapping Using Text Mining: A Foresight Study for the Retail Industry. IEEE Transactions on Engineering Management 69, 228–244. https://doi.org/10.1109/TEM.2021.3068310

Park, H., Phaal, R., Ho, J.-Y., O'Sullivan, E., 2020. Twenty years of technology and strategic roadmapping research: A school of thought perspective. Technological Forecasting and Social Change 154, 119965. https://doi.org/10.1016/j.techfore.2020.119965

Peng, Y., Tan, A., Wu, J., Bi, Y., 2019. Hierarchical Edge Computing: A Novel Multi-Source Multi-Dimensional Data Anomaly Detection Scheme for Industrial Internet of Things. IEEE Access 7, 111257–111270. https://doi.org/10.1109/ACCESS.2019.2930627

Phaal, R., Farrukh, C., Probert, D., 2004a. Customizing Roadmapping. Research-Technology Management 47, 26–37. https://doi.org/10.1080/08956308.2004.11671616

Phaal, R., Farrukh, C., Probert, D., 2001. Technology Roadmapping: Linking Technology Resources to Business Objectives. Univ. Cambridge 1–18.

Phaal, R., Farrukh, C.J.P., Probert, D.R., 2004b. Technology roadmapping—A planning framework for evolution and revolution. Technological Forecasting and Social Change, Roadmapping: From Sustainable to Disruptive Technologies 71, 5–26. https://doi.org/10.1016/S0040-1625(03)00072-6

Phaal, R., O'Sullivan, E., Routley, M., Ford, S., Probert, D., 2011. A framework for mapping industrial emergence. Technological Forecasting and Social Change, Using Technological Intelligence for Strategic Decision Making in High Technology Environments 78, 217–230. https://doi.org/10.1016/j.techfore.2010.06.018

Pop, P., Zarrin, B., Barzegaran, M., Schulte, S., Punnekkat, S., Ruh, J., Steiner, W., 2021. The FORA Fog Computing Platform for Industrial IoT. Information Systems 98, 101727. https://doi.org/10.1016/j.is.2021.101727

Pora, U., Gerdsri, N., Thawesaengskulthai, N., Triukose, S., 2022. Data-Driven Roadmapping (DDRM): Approach and Case Demonstration. IEEE Trans. Eng. Manage. 69, 209–227. https://doi.org/10.1109/TEM.2020.3005341

Portal, G., de Matos, E., Hessel, F., 2020. An Edge Decentralized Security Architecture for Industrial IoT Applications, in: 2020 IEEE 6th World Forum on Internet of Things (WF-IoT). Presented at the 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), pp. 1–6. https://doi.org/10.1109/WF-IoT48130.2020.9221176

Pousttchi, K., Gleiss, A., Buzzi, B., Kohlhagen, M., 2019. Technology Impact Types for Digital Transformation, in: 2019 IEEE 21st Conference on Business Informatics (CBI). Presented at the 2019 IEEE 21st Conference on Business Informatics (CBI), IEEE, Moscow, Russia, pp. 487–494. https://doi.org/10.1109/CBI.2019.00063

Qiu, T., Chi, J., Zhou, X., Ning, Z., Atiquzzaman, M., Wu, D.O., 2020. Edge Computing in Industrial Internet of Things: Architecture, Advances and Challenges. IEEE Communications Surveys Tutorials 22, 2462–2488. https://doi.org/10.1109/COMST.2020.3009103

Reimers, N., Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. https://doi.org/10.48550/arXiv.1908.10084

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H.R., Albarqouni, S., Bakas, S., Galtier, M.N., Landman, B.A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R.M., Trask, A., Xu, D., Baust, M., Cardoso, M.J., 2020. The future of digital health with federated learning. npj Digit. Med. 3, 1–7. https://doi.org/10.1038/s41746-020-00323-1

Rush, H., Bessant, J., Hobday, M., 2007. Assessing the technological capabilities of firms: developing a policy tool. R&D Management 37, 221–236. https://doi.org/10.1111/j.1467-9310.2007.00471.x

Satyanarayanan, M., 2017. The Emergence of Edge Computing. Computer 50, 30–39. https://doi.org/10.1109/MC.2017.9

Satyanarayanan, M., Bahl, P., Caceres, R., Davies, N., 2009. The Case for VM-Based Cloudlets in Mobile Computing. IEEE Pervasive Computing 8, 14–23. https://doi.org/10.1109/MPRV.2009.82

Shaik, M.S., Struhár, V., Bakhshi, Z., Dao, V.-L., Desai, N., Papadopoulos, A.V., Nolte, T., Karagiannis, V., Schulte, S., Venito, A., Fohler, G., 2020. Enabling Fog-based Industrial Robotics Systems, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Presented at the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 61–68. https://doi.org/10.1109/ETFA46521.2020.9211887

Sievert, C., Shirley, K., 2014. LDAvis: A method for visualizing and interpreting topics, in: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces. Presented at the Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Association for Computational Linguistics, Baltimore, Maryland, USA, pp. 63–70. https://doi.org/10.3115/v1/W14-3110

Sittón-Candanedo, I., Alonso, R.S., Corchado, J.M., Rodríguez-González, S., Casado-Vara, R., 2019. A review of edge computing reference architectures and a new global edge proposal. Future Generation Computer Systems 99, 278–294. https://doi.org/10.1016/j.future.2019.04.016

Son, W., Lee, S., 2019. Integrating fuzzy-set theory into technology roadmap development to support decision-making. Technology Analysis & Strategic Management 31, 447–461. https://doi.org/10.1080/09537325.2018.1522432

StarlingX, 2021. Open Source Edge Cloud Computing Architecture - StarlingX [WWW Document]. URL https://starlingx.io/ (accessed 12.20.21).

Syed, S., Spruit, M., 2017. Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation, in: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). Presented at the 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 165–174. https://doi.org/10.1109/DSAA.2017.61

Taïk, A., Cherkaoui, S., 2020. Electrical Load Forecasting Using Edge Computing and Federated Learning, in: ICC 2020 - 2020 IEEE International Conference on Communications (ICC). Presented at the ICC 2020 - 2020 IEEE International Conference on Communications (ICC), pp. 1–6. https://doi.org/10.1109/ICC40277.2020.9148937

The Linux Foundation, 2021. Why EdgeX [WWW Document]. https://www.edgexfoundry.org/why_edgex/why-edgex/. URL https://www.edgexfoundry.org/why_edgex/why-edgex/ (accessed 12.20.21).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is All you Need, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Vial, G., 2019. Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, SI: Review issue 28, 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Wang, B., Liu, Y., Zhou, Y., Wen, Z., 2018. Emerging nanogenerator technology in China: A review and forecast using integrating bibliometrics, patent analysis and technology roadmapping methods. Nano Energy 46, 322–330. https://doi.org/10.1016/j.nanoen.2018.02.020

Willyard, C.H., McClees, C.W., 1987. Motorola's Technology Roadmap Process. Research Management 30, 13–19. https://doi.org/10.1080/00345334.1987.11757057

Wu, D., Liu, S., Zhang, L., Terpenny, J., Gao, R.X., Kurfess, T., Guzzo, J.A., 2017. A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing. Journal of Manufacturing Systems 43, 25–34. https://doi.org/10.1016/j.jmsy.2017.02.011

Wu, Y., Dai, H.-N., Wang, H., 2021. Convergence of Blockchain and Edge Computing for Secure and Scalable IIoT Critical Infrastructures in Industry 4.0. IEEE Internet Things J. 8, 2300–2317. https://doi.org/10.1109/JIOT.2020.3025916

Xu, W., Zhou, H., Cheng, N., Lyu, F., Shi, W., Chen, J., Shen, X., 2018. Internet of vehicles in big data era. IEEE/CAA Journal of Automatica Sinica 5, 19–35. https://doi.org/10.1109/JAS.2017.7510736

Xu, Z., Zhao, L., Liang, W., Rana, O.F., Zhou, P., Xia, Q., Xu, W., Wu, G., 2021. Energy-Aware Inference Offloading for DNN-Driven Applications in Mobile Edge Clouds. IEEE Transactions on Parallel and Distributed Systems 32, 799–814. https://doi.org/10.1109/TPDS.2020.3032443

Yi, S., Hao, Z., Qin, Z., Li, Q., 2015. Fog Computing: Platform and Applications, in: 2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb). Presented at the 2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb), pp. 73–78. https://doi.org/10.1109/HotWeb.2015.22

Yu, W., Liang, F., He, X., Hatcher, W.G., Lu, C., Lin, J., Yang, X., 2018. A Survey on the Edge Computing for the Internet of Things. IEEE Access 6, 6900–6919. https://doi.org/10.1109/ACCESS.2017.2778504

Zaki, M., 2019. Digital transformation: harnessing digital technologies for the next generation of services. Journal of Services Marketing 33, 429–435. https://doi.org/10.1108/JSM-01-2019-0034

Zhang, H., Daim, T., Zhang, Y. (Peggy), 2021. Integrating patent analysis into technology roadmapping: A latent dirichlet allocation based technology assessment and roadmapping in the field of Blockchain. Technological Forecasting and Social Change 167, 120729. https://doi.org/10.1016/j.techfore.2021.120729

Zhang, Y., Guo, Y., Wang, X., Zhu, D., Porter, A.L., 2013. A hybrid visualisation model for technology roadmapping: bibliometrics, qualitative methodology and empirical study. Technology Analysis & Strategic Management 25, 707–724. https://doi.org/10.1080/09537325.2013.803064

Zhang, Y., Zhang, G., Chen, H., Porter, A.L., Zhu, D., Lu, J., 2016. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. Technological Forecasting and Social Change 105, 179–191. https://doi.org/10.1016/j.techfore.2016.01.015

Zhang, Y., Zhou, X., Porter, A.L., Vicente Gomila, J.M., 2014. How to combine term clumping and technology roadmapping for newly emerging science & technology competitive intelligence: "problem & solution" pattern based semantic TRIZ tool and case study. Scientometrics 101, 1375–1389. https://doi.org/10.1007/s11192-014-1262-2