

Sara Hiljanen

# HOW IS EDGE COMPUTING AFFECTING REMOTE SERVICES IN THE PROCESS INDUSTRY?

Risks, opportunities, and required capabilities

Master of Science Thesis  
Faculty of  
Management and Business  
Examiner: Hannu Kärkkäinen  
Examiner: Henri Pirkkalainen  
June 2024

# ABSTRACT

Sara Hiljanen: How edge computing is affecting remote services in the process industry? – Risks, opportunities, and required capabilities

Master of Science Thesis

Tampere University

Master's Degree Programme in Information and Knowledge Management

June 2024

---

The process industry has moved towards digital servitization and data-driven services to provide value for the end customer. A common way has been to do remote monitoring in the cloud and perform remote services with cloud-based IoT applications and other analytical capabilities. Since the amount of data is increasing and the requirements for remote services are changing, the cloud can not answer all demands anymore. The huge amount of data has caused latency and increased data processing costs in the cloud. Hence a new solution edge has been proposed to answer the real-time requirements and to reduce the burden of the cloud for a better user experience and minimized latency and costs.

The edge solution is still a new solution, which is why there is yet no clear definition of how it should be utilized in the process industry and what are its roles compared to the cloud. This research was done as a case study. The case company wanted to understand how an edge solution will change the traditional way of performing remote services in the cloud and what kind of capabilities are needed to keep performing remote services at the edge. Hence, this research aims to understand how edge solution will change how remote services are performed, and what needs to be taken into account when developing edge solution for enabling remote services at the edge. The base understanding and base theory of the topic were collected from the case company and with a literature review. The defined theory was then tested with a case study in the case company consisting of a pre-questionnaire and interviews with research participants to collect both qualitative and quantitative data.

As a result, findings from the literature review and case study were combined to answer the research problem and seven sub-research questions. It was found that the edge cloud can not fully replace the cloud since the cloud has still some strengths and opportunities that the edge can not meet like fleet monitoring and training more advanced analytical models. Edge solution can broaden the service and product portfolio by enabling local solutions of the offering for customers who do not want to transfer their data to the cloud. Edge can also act as an extension of the cloud to answer real-time requirements and lower the burden of cloud computing and possibly minimize costs. This model is a hybrid model where the strengths of both cloud and edge are utilized to achieve the best outcome. In the hybrid model, the edge can handle pre-processing data that can then be transferred to the cloud for post-processing. Cloud can handle training more advanced analytical models which edge does not have enough capacity for and the trained model attributes can be transferred to edge for local implementation. To enable performing remote services at the edge, there needs to be remote access which enables easy access to the solution. Additionally, the roles of edge and cloud need to be clearly defined in a hybrid model.

This research guides companies working in the process industry and providing remote services about how edge solution should be utilized and what needs to be considered when developing the solution. Additionally, the research creates a base for further studies. The topic is wide which is why there are a lot of possibilities for future work. For literature, this research fills the identified gap and provides a further understanding of the edge in the context of process industry and remote services.

Keywords: Process industry, IoT, data-driven services, remote services, digital servitization, edge computing, cloud computing

The originality of this thesis has been checked using the Turnitin Originality Check service.

# TIIVISTELMÄ

Sara Hiljanen: Kuinka reunalaskenta vaikuttaa etäpalveluiden tarjoamiseen prosessiteollisuudessa? – Uhat, mahdollisuudet ja vaaditut kyvykkyydet

Diplomityö

Tampereen yliopisto

Tietojohtamisen diplomi-insinöörin tutkinto-ohjelma

Kesäkuu 2024

Digitaalinen palvelullistaminen ja datapohjaiset palvelut ovat yleistyneet prosessiteollisuudessa arvontuottamiseksi asiakkaille. Etäseurantaa on yleensä tarjottu pilvessä pilvipohjaisten esineiden internetin sovellusten ja muiden analyysikyvykkyyden avulla. Pilvi ei kuitenkaan pysty vastaamaan datan määrän kasvun seurauksena muuttuviin etäseurannan tarpeisiin. Suuri datan määrä on aiheuttanut viivettä ja kohonneita kustannuksia pilveen. Tämän takia uudenlaista reunalaskenta ratkaisua on ehdotettu vähentämään pilven kuormitusta taatakseen paremman käyttäjäkokemuksen, pienentämään viivettä ja mahdollisesti vähentämään kustannuksia.

Reunalaskenta on uusi ratkaisu, jonka takia vielä ei ole selkeää kuvausta kuinka sitä tulisi hyödyntää prosessiteollisuudessa tai mikä sen rooli on pilveen verrattuna. Tämä tutkimus tehtiin tapaustutkimuksena. Kohdeyritys halusi ymmärtää kuinka reunalaskenta muuttaa etäpalveluiden toteuttamista sekä millaisia kyvykkyyksiä tarvitaan, jotta etäpalveluita pystytään toteuttamaan reunalaskennan avulla. Siksi tämän tutkimuksen tavoite oli tutkia kuinka reunalaskenta muuttaa tapaa toteuttaa etäpalveluita ja millaisia kyvykkyyksiä tarvitaan etäpalveluiden toteuttamiseksi reunalaskennan avulla. Pohjatiedot aiheesta kerättiin kohdeyritykseltä, ja alustava teoria muodostettiin pohjatiedon ja kirjallisuuden pohjalta. Tätä teoriaa testattiin tapaustutkimuksella kohdeyrityksessä kyselyiden ja haastatteluiden avulla. Kyselyissä ja haastatteluissa kerättiin sekä kvalitatiivista että kvantitatiivista dataa.

Löydökset kirjallisuudesta ja tapaustutkimuksesta muodostivat vastauksen tutkimusongelmaan ja sen alatutkimuskysymyksiin. Löydöksen perusteella reunalaskenta ei pysty korvaamaan pilvilaskentaa täysin, sillä pilvellä on vahvuuksia ja kehitysmahdollisuuksia, joita reunalaskenta ei mahdollista. Näitä vahvuuksia ovat esimerkiksi fleetin seuranta ja älykkäiden analytiikkamallien kouluttaminen. Reunalaskenta laajentaa palvelu- ja tuoteportfoliota mahdollistaen paikallisten ratkaisuiden toteuttamisen asiakkaille, jotka eivät halua lähettää dataansa pilveen. Reunalaskenta voi toimia myös pilvilaskennan jatkeena vastatakseen reaaliaikaisuuden tarpeeseen, helpottamaan pilvilaskennan kuormitusta sekä mahdollisesti vähentämään kustannuksia. Tällaista toimintamallia kutsutaan hybridimalliksi, jossa hyödynnetään reunalaskennan ja pilvilaskennan parhaita ominaisuuksia optimaalisen lopputuloksen saavuttamiseksi. Hybridimallissa reunalaskennan avulla voidaan prosessoida dataa, joka sitten siirretään pilveen jatkokäsittelyä varten. Pilvilaskennan avulla voidaan kouluttaa älykkäitä analytiikkamalleja, joihin reunalaskennassa ei ole kapasiteettia. Mallien parametrit voidaan lähettää reunalaskentaan paikallista toteutusta varten. Jotta etäpalveluita voidaan toteuttaa reunalaskennan avulla, tulee ratkaisuihin päästä helposti käsiksi etäyhteyden avulla. Lisäksi reunalaskennan ja pilvilaskennan roolit tulee määritellä hybridimallissa.

Tämä tutkimus tarjoaa ohjeistusta etäpalveluita prosessiteollisuudessa tarjoaville yrityksille reunalaskennan hyödyntämiseen sekä siihen, millaisia asioita pitää huomioida reunalaskennan kehityksessä. Lisäksi tämä tutkimus tarjoaa teoreettisen pohjan jatkotutkimukselle. Aihe on laaja, minkä takia aiheen parissa on paljon potentiaalia ja tarvetta jatkotutkimukselle. Kirjallisuuden näkökulmasta tämä työ täyttää tunnistettuja puutteita ja tarjoaa ymmärrystä reunalaskennasta prosessiteollisuuden ja etäpalveluiden kontekstissa.

Avainsanat: Prosessiteollisuus, esineiden internet, datapohjaiset palvelut, etäpalvelut, digitaalinen palvelullistaminen, reunalaskenta, pilvilaskenta

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

## PREFACE

I'm grateful that I was able to do this thesis on such an interesting and topical topic. This is a good way to finish my five years of studies with Information and Knowledge Management which have taught me a lot. However, I must say that most of my learning has happened at work by experiencing and doing.

Thank you to Hannu Kärkkäinen for your spectacular guidance and bits of advice during these last six months. Thank you to Tuomas, Jari, and Heikki from the case company side for this opportunity, guidance, and the knowledge sharing that we have gone through during the work. Special thanks to Otto for sparring and supporting me during this work. I want to also thank all the people who participated in this research and made this possible.

I want to provide my warmest thanks to Samu, Sonja, and my family for the tremendous support you have given me, it motivated me to get this thesis done even earlier than planned. Thank you for caring for my well-being during this time and always.

Helsinki, 19 June 2024

Sara Hiljanen

# CONTENTS

1. INTRODUCTION .....	1
1.1 Background.....	1
1.2 Research objectives and questions.....	1
1.3 Research focus, boundaries, and structure .....	3
2. RESEARCH DESIGN AND METHODS .....	5
2.1 Research design.....	5
2.2 Method for literature review.....	7
2.3 Method for Case Study .....	10
3. PROCESS INDUSTRY AND SERVICITIZATION .....	16
3.1 Process Industry and its characteristics .....	16
3.1.1 Different ways to do analytics and gain business value.....	17
3.2 Challenges in Process Industry.....	20
3.3 Services in process industry.....	20
4. DIGITAL SERVICITIZATION IN PROCESS INDUSTRY .....	23
4.1 Digital servicitization and smart PSS .....	23
4.2 Data-driven services .....	24
4.2.1 Challenges in data-driven services .....	27
4.3 Remote services with cloud computing .....	27
5. REMOTE SERVICES WITH EDGE COMPUTING .....	30
5.1 Edge computing.....	30
5.2 Edge cloud hybrid model.....	32
5.2.1 Different layers of the hybrid model.....	33
5.2.2 Placement decision in hybrid model.....	35
5.3 Cloud, edge, or both?.....	37
6. RESULTS .....	40
6.1 Pre-questionnaire.....	40
6.1.1 Maturity level model questions .....	43
6.2 Interview results .....	44
6.2.1 Single cloud strengths, weaknesses, opportunities, and threats..	45
6.2.2 Single edge strengths, weaknesses, opportunities and threats ...	47
6.2.3 DCS strengths, weaknesses, opportunities and threats .....	49
6.2.4 Critical capabilities for remote services .....	50
6.2.5 Different criticality of requirements inside the case company .....	54
7. DISCUSSION.....	57
7.1 Remote services in process industry .....	57
7.2 Tools and capabilities for performing remote services .....	61
7.3 Challenges of cloud.....	64
7.4 Risks and opportunities of edge for remote services .....	65

7.5	Role of edge in remote services .....	67
7.6	Ways to ensure performing remote services with edge solution .....	69
8.	CONCLUSIONS.....	73
8.1	Answers to research questions .....	73
8.2	Managerial implications.....	75
8.2.1	Overall managerial recommendations.....	76
8.3	Evaluation of research .....	77
8.4	Future work.....	79
	REFERENCES.....	80

## LIST OF FIGURES

<b>Figure 1.</b>	<i>Research onion model (adapted from Saunders et al. 2007)</i> .....	5
<b>Figure 2.</b>	<i>Process for information search (adapted from Järvelin &amp; Sormunen 2010)</i> .....	7
<b>Figure 3.</b>	<i>The relationships of materials and the most remarkable materials utilized in this research</i> .....	10
<b>Figure 4.</b>	<i>Business value in proactive and reactive actions (adapted from Krumeich et al. 2016)</i> .....	18
<b>Figure 5.</b>	<i>Analytics maturity model (adapted from Halper 2020; Wloniak &amp; Grebski 2023)</i> .....	19
<b>Figure 6.</b>	<i>From meaningful data to wisdom (adapted from Ackoff 1989; Rapaccini &amp; Adrodegari 2022; Zheng et al. 2019)</i> .....	25
<b>Figure 7.</b>	<i>Service types based on who is conducting the service (adapted from Rapaccini &amp; Adrodegari 2022)</i> .....	26
<b>Figure 8.</b>	<i>Automation hierarchy and real-time requirements of each hierarchy level (adapted from Dai et al. 2023; Jin et al. 2024)</i> . ....	28
<b>Figure 9.</b>	<i>area coverage and respnse time in automation systems, edge and cloud (adapted from Mubeen et al. 2017)</i> .....	31
<b>Figure 10.</b>	<i>Edge and cloud potential in automation hierarchy (adapted from Jin et al. 2024)</i> .....	33
<b>Figure 11.</b>	<i>Edge-cloud-integration model (adapted from Liu et al. 2021)</i> .....	34
<b>Figure 12.</b>	<i>Framework for placement decisions in cloud, edge, and thing levels (adapted from Hästbacka et al. 2022)</i> .....	36
<b>Figure 13.</b>	<i>Comparison between single edge, single cloud, and a hybrid model</i> .....	37
<b>Figure 14.</b>	<i>The maturity level of analytics in the case company</i> .....	44
<b>Figure 15.</b>	<i>Capabilities required for remote services</i> . ....	52
<b>Figure 16.</b>	<i>Traditional way-of-working</i> .....	59
<b>Figure 17.</b>	<i>Proactive way-of-working</i> .....	60
<b>Figure 18.</b>	<i>SWOT analysis of single cloud, single edge, hybrid solutions and DCS</i> .....	67
<b>Figure 19.</b>	<i>Task division between cloud, edge, and DCS</i> .....	68

## LIST OF TABLES

<i>Table 1.</i>	<i>Search inquiries and found materials .....</i>	<i>8</i>
<i>Table 2.</i>	<i>Research participant information.....</i>	<i>11</i>
<i>Table 3.</i>	<i>Questions in pre-questionnaire .....</i>	<i>12</i>
<i>Table 4.</i>	<i>Development plans for remote services in the case company.....</i>	<i>41</i>
<i>Table 5.</i>	<i>Single cloud strengths, weaknesses, opportunities, and threats.....</i>	<i>45</i>
<i>Table 6.</i>	<i>Single edge strengths, weaknesses, opportunities, and threats .....</i>	<i>47</i>
<i>Table 7.</i>	<i>DCS strengths, weaknesses, opportunities, and threats .....</i>	<i>50</i>
<i>Table 8.</i>	<i>Required tools and capabilities for performing remote services.....</i>	<i>61</i>
<i>Table 9.</i>	<i>Things that need to be defined when developing an edge solution in the case company.....</i>	<i>70</i>

# LIST OF SYMBOLS AND ABBREVIATIONS

DCS	Distributed Control System
IoT	Internet of Things
IIoT	Industrial Internet of Things
OEM	Original Equipment Manufacturer
AI	Artificial Intelligence
OEE	Overall Equipment Effectiveness
SaaS	Software as a Service
KPI	Key Performance Indicator

# 1. INTRODUCTION

This chapter introduces the background of the topic, research objectives and questions, and research focus, boundaries, and structure.

## 1.1 Background

One main challenge in the process industry is to reach continuous 100% availability (Isaksson et al. 2018). Hence continuous monitoring and optimization of processes is a must in the process industry to stay competitive and minimize unplanned downtime. (Azeved 2022; Krumeich et al. 2016) Since technology is developing all the time and becoming more complex, another challenge in the process industry is how to maintain these more complex systems with less effort (Isaksson et al. 2018). As the systems get more complex, IoT has also enabled collecting of large amounts of data from the monitored assets (Kaur et al. 2018). This has created opportunities for new kinds of services with digital capabilities, data-driven services. Collected data itself is meaningless unless it is processed, analyzed, evaluated, and utilized in decision-making. (Rapaccini & Adrodegari 2022) Cloud computing is a common way of performing remote services in the process industry since it enables more efficient monitoring and more intelligent analytics that can help in monitoring tasks (Raddats et al. 2022; Ren et al. 2022.).

Even though cloud computing has been proven to be an efficient way for efficient way for processing data, it has limitations when the amount of data is increasing (Shi et al. 2016). It can not answer the real-time requirements since it consists of latency (Liu et al. 2019). Hence a new solution, edge computing has been proposed to support the cloud in these kinds of requirements and to minimize the burden of the cloud (Kong et al. 2022; Shi et al. 2016). However, it is not straightforward to determine how edge computing should be utilized in the process industry and how is it positioned with cloud computing (Ashouri et al. 2018).

## 1.2 Research objectives and questions

This research concentrates on an edge solution and how it changes the current way of conducting remote services. The need for this research arose from a case company utilized in this research. The case company is a large Finnish process technologies and

automation company that performs remote services for its customers. The case company is developing an edge solution in addition to their current cloud solution. The drivers for this development have been that some customers are unwilling to transfer their data to the cloud but are still interested in the services and products offered in the cloud. Another driver for the edge is data processing costs. The hypothesis in the case company is that an edge solution could help to decrease the data processing costs. The third driver for the edge solution is that there are IoT applications that should work based on real-time data but since they are residing in the cloud, they contain latency. The second hypothesis in the case company is that the cloud is no longer needed but the edge solution could be utilized instead. This research aims to address these hypotheses and investigate how edge solution could change the way of performing remote services and hence what kind of things should be taken into account when developing the edge solution further to enable performing remote services with edge solution. Based on El-Sayed et al. (2017) scalability, robustness, efficiency, manageability, and dynamicity are not usually in focus when developing an edge solution. Hence this research helps to close this gap and provide a better understanding and guidance on how these aspects can be better taken into account when developing an edge solution. There is no existing literature about how edge solution changes the way of performing remote services which is critical to consider in the process industry. This research aims to close this gap. Based on these targets, the main research question and its sub-research questions are:

**Main RQ:** How edge solution will change the way of performing remote services, and what kinds of things are critical for enabling performing remote services when developing an edge solution?

**Sub RQ1:** What kind of remote services are provided in the process industry?

**Sub RQ2:** What kind of technical capabilities and tools are required for conducting remote services?

**Sub RQ3:** What kind of shortages there is currently in capabilities in the cloud?

**Sub RQ4:** What kind of benefits edge solution could provide for remote services compared to a single cloud solution?

**Sub RQ5:** What kind of risks and challenges there is in single edge solution?

**Sub RQ6:** How edge solution change the way remote services are performed?

**Sub RQ7:** Which things should be defined in the case company when developing edge solution further?

### **1.3 Research focus, boundaries, and structure**

Since this research is conducted as a case study, the focus is on the process industry, Industrial Internet, and remote services. The remote services are narrowed down to services that are currently conducted in the cloud since those are the services and products that will be at the edge in the future in the case company. The products indicate IoT applications and technical capabilities. In the case company, the edge solution is a server that contains the capability to hold cloud-based IoT applications and technical capabilities that a service provider can utilize to perform remote services. The customer has also access to the edge and its products and capabilities. In this research, the edge solution is studied without narrowing down the focus to only this kind of edge solution to get a wider understanding of the possible capabilities. This research will focus more on the way of working and the business side of the topic rather than technical specifications. This research is a base for further research related to the topic which can include more technical specifications. If both would be included in this research, the scope would be too wide.

This research is divided into a theoretical foundation part consisting of a literature review and an empirical analysis part consisting of a case study. The literature review is presented in chapters 3-5. The chapter 3 presents the process industry and its main characteristics to create a base understanding of the context and to understand which characteristics of the process industry affect this research. In chapter 3 also the analytics maturity model is created that is used to analyze the maturity level of the analytics in the case company.

Since chapter 3 does not yet take into account how remote services are performed in the process industry, chapter 4 presents the current way of performing remote services in the process industry with digital capabilities. Currently, the traditional way for conducting remote services in the process industry has been to utilize cloud capabilities together with traditional DCS. Chapter 4 presents also the challenges in the current way of performing remote services in the process industry which creates a base understanding of why the edge is needed.

Finally, chapter 5 presents edge solution and the opportunities that it can bring to the way of performing remote services in the process industry. Since single single-edge solution has its restrictions, a hybrid model is studied as a third option. Chapter 5 aims to compare a single-cloud, a single-edge, and a hybrid model.

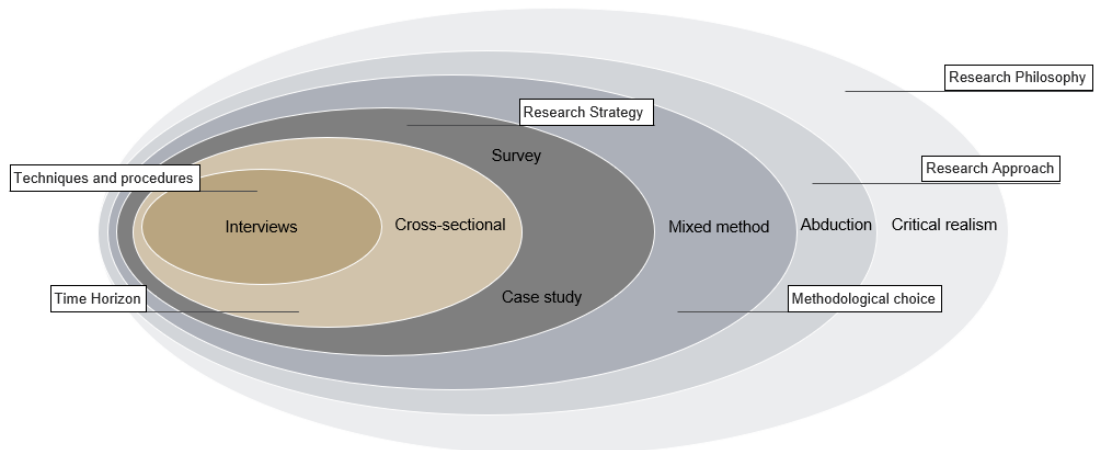
The results of the case study are presented in chapter 6 where also the maturity level of the analytics in the case company is defined, and the findings and discussion are presented in chapter 7. Chapter 7 answers the research questions and highlights the most interesting findings. Finally, a conclusion containing short answers to the research questions, managerial implications, research evaluation, and future work is presented in chapter 8. First chapter 2 presents the research design and methods.

## 2. RESEARCH DESIGN AND METHODS

This chapter presents the decisions that were made when designing this research by utilizing the research onion model from Saunders et al. (2007). The other half of this chapter presents the methods for literature review that were used to create a base theory for the research and methods for the empirical part of the research.

### 2.1 Research design

The research onion model from Saunders et al. (2007) was utilized as a guideline when designing this research. The research onion model was chosen since it gives clear guidelines for the required decisions and makes the decisions transparent. Usually when research is designed, the techniques are the first thing that are considered for answering research questions. However, the techniques require multiple choices that result in the decision. These decisions need to be made visible so that the research can be trusted. (Saunders et al. 2007) The onion model and the decisions made when designing this research are illustrated in Figure 1.



**Figure 1.** Research onion model (adapted from Saunders et al. 2007)

The outer layer in the onion model is **research philosophy**. Research philosophy refers to beliefs and assumptions. The assumptions can refer to assumptions about the nature of reality or about knowledge. The beliefs refer to values and ethics. (Saunders et al. 2007) This research aims to identify how edge computing can affect remote services in the process industry but to also guide what things should be defined and developed further related to edge cloud so that it does not put remote services at risk. This is the research problem that is answered in this research and hence pragmatism is suitable as a research philosophy in this research since in pragmatism the aim is to follow a research

problem and the focus is on practical solutions and outcomes. Pragmatism focuses on practical solutions and outcomes to make an actual difference in organizational practices. Pragmatism sees that the research is only relevant if it supports action. (Saunders et al. 2007)

The second layer from outside the onion model is the **research approach**. The research approach is related to theory, whether it is tested or created through research. In deduction, the existing theory is tested through hypotheses. As a result, the existing theory can be falsified or verified. In induction, research explores a phenomenon, identifies themes and patterns, and creates a conceptual framework that can then be generalized into a new theory. Abduction differentiates from deduction and induction so that it utilizes both, existing theory, and new data to create a generalization. Abduction moves back and forth between theory and data to modify or create theory. (Saunders et al. 2007) In this research, the theory is utilized as a base for the research. After the research, the results are compared to the theory to create a new modified theory and framework. Hence, this research takes partial abduction as a research approach.

The third layer from the outside of the onion model is **methodological choices**. The first methodological choice is whether the research utilizes qualitative, quantitative, or mixed methods. The qualitative method generates non-numerical data through data collection techniques like interviews. The quantitative method generates instead numeric data through data collection techniques like surveys. If the research combined both, non-numerical data and numeric data, the research utilizes mixed method. (Saunders et al. 2007) This research utilizes both, qualitative and quantitative data collection, to provide a broader understanding of the research problem and hypotheses and hence the mixed method is chosen.

The next layer towards the center of the onion model is the **research strategy**. A simple definition for research strategy is a plan of actions to achieve the defined goals. The choice of whether qualitative, quantitative, or mixed method is used affects the research strategy since the research strategy points out the concrete method for data collection. There are many options to choose from which interviews and surveys were already mentioned. (Saunders et al. 2007) This research utilizes interviews to collect non-numerical data and surveys to collect numeric data. However, the interviews and surveys are done for a case company which narrows the scope to a specific real-life setting hence case study is one research strategy in addition to the survey.

The second layer from the middle of the onion model is the **time horizon**. The time horizon can be either cross-sectional where a phenomenon is studied at a particular time

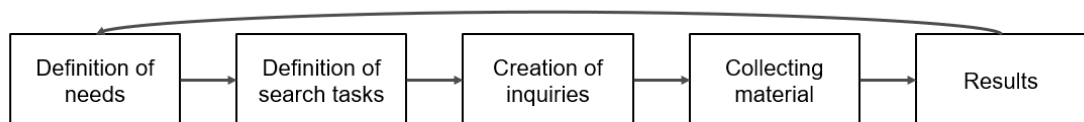
or longitudinal where the phenomenon is studied over a longer time. Longitudinal studies enable recognizing change and development. (Saunders et al. 2007) This research is conducted during six months and hence the time horizon is cross-sectional.

The midmost layer in the onion model is the **techniques and procedures** for data collection and analysis (Saunders et al. 2007). As said earlier, interviews and surveys are utilized for the data collection. In the analysis, the numeric data is grouped for themes that are then combined with non-numeric data to create the results of the research.

This research combines multiple methods to create actionable outcomes that can create value for an organization when applied. This supports the definition of pragmatism that was chosen as a research philosophy. The decision to utilize both qualitative and quantitative data collection resulted in the decision to utilize case study and surveys as a research strategy, and for choosing surveys and interviews as techniques for data collection. Since the time horizon is cross-sectional, this research does not study change or development but instead time-bound situation. Since this study utilizes the abduction approach, the aim is to either modify or create a theory in the end. Even though the research is a case study, the abduction approach enables the creation of results that can be generalized. Cross-sectional time horizon however makes the results time-bound and the results might need to be validated after time passes. The research onion model was not mechanically followed but instead utilized as a guideline in the design of this research and for demonstrating the decisions that were made.

## 2.2 Method for literature review

Since the internet and search engines have developed, information search has become a routine for everyone. Information search aims to find relevant information for those who need it. The search process goes as presented in Figure 2. First needs need to be defined. After that related search task is formed and suitable search inquiries are defined. (Järvelin & Sormunen 2010)



**Figure 2.** Process for information search (adapted from Järvelin & Sormunen 2010)

In this research, the need was to find information about how edge computing will change the way remote services are performed and hence the search task was to find information about the process industry, how remote services are conducted in the process

industry, cloud computing, what kind of challenges there is when performing remote services, and edge computing. The suitable search inquiries are presented in Table 1.

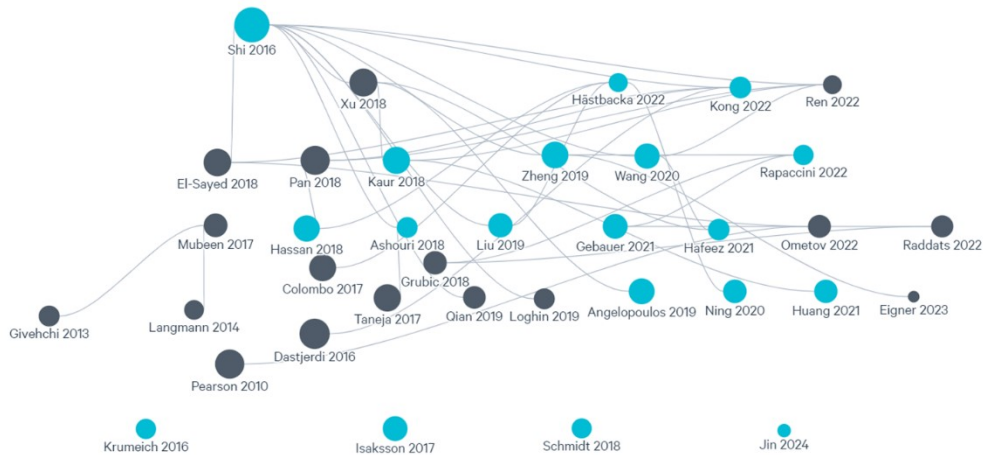
Table 1. Search inquiries and found materials

<b>Source</b>	<b>Search phrase</b>	<b>Results</b>	<b>Chosen articles</b>
<i>Andor</i>	"data-driven services" AND industrial	93 items	Rapaccini & Adrodegari 2022
	edge AND "automation system" AND indust*	481 items	Mubeen et al. 2017, Dai et al. 2023, Jin et al. 2024
	"process automation" AND edge AND cloud	89 items	Isaksson et al. 2018, Xu et al. 2018, Tusa et al. 2024, Angelopoulos et al. 2019
<i>Scopus</i>	"process industry" AND ( ( ( "data-driven" OR smart ) AND service* ) OR servitiza* )	32 documents	Schmidt et al. 2018, Kanninen et al. 2017
<i>Google Scholar</i>	"edge" AND "smart product service"	965 documents	Liu et al. 2019, Zheng et al. 2019, Hästbacka et al. 2021
	edge AND "predictive maintenance"	30k documents	Hafeez et al. 2021, Teoh et al. 2021, Bellavista et al. 2020
	"edge computing" AND ("challenges"	222k documents	Shi et al. 2016, Kaur et al. 2018,

OR "industrial internet of things")		Hassan et al. 2018, Pan et al. 2017
"cloud" AND "security" AND "edge"	1m documents	Ometov et al. 2022, El-Sayed et al. 2017, Francis et al. 2018, Ashouri et al. 2018

With these search phrases, in Table 1, the first sources were found as listed in the table. After this the articles were studied and new needs related to materials were defined (Järvelin & Sormunen 2010). This kind of approach is called citation pearl growing which means that when a relevant material is found, it is studied and new terms are collected that can be used to broaden up the search inquiries (Booth 2015). Then the search process started again. To be able to critically evaluate the found materials, there need to be evaluation criteria (Järvelin & Sormunen 2010). In this research, the found material needs to be of course relevant for the research, scientific, and consist of up-to-date information. The materials utilized in this research were mainly done during the last ten years but some older materials were used if the information they contained was still up-to-date. Thesis works and websites were not utilized except if it was possible to detect that the material was scientific. For example, the material by Halper (2020) was accepted since it had been cited in a scientific paper.

In addition to the previously presented method, the backward and forward snowballing methods were used. Both methods start by identifying a starting set of materials. In backward snowballing, the references of the existing material are examined to find new potential materials to add to the existing set. In forward snowballing the starting point is the same but instead of examining references, citations are examined to identify materials that have referenced the examined material. (Wohlin 2014) Figure 3 illustrates the connections between different materials utilized in this research. The most remarkable materials for this research are highlighted in blue.



**Figure 3.** *The relationships of materials and the most remarkable materials utilized in this research.*

It can be seen that not all materials in this research are connected. This occurs because citation pearl growing was utilized which does not require that the new materials found with the existing material are connected. The existing material can for example give improvement ideas for the search inquiries and result in finding the new material. It can be noticed that still many materials in this research are connected which is due to the backward and forward snowballing effect.

### 2.3 Method for Case Study

This chapter presents the case company and the case study. The case company is a leading supplier for process technologies, automation, and services globally. Hence the research participants were chosen from all three businesses. The case company operates in the pulp, paper, and energy industries but the industries were not separated in this research. The offering of the case company consists of machinery, equipment, and systems. In this research, the focus is on the Industrial Internet offering of the company. Industrial Internet combines expertise in the case company in process technology, automation, and services with data analytics for improved profitability of their customers and more autonomous operations. This can be done by maximizing asset and people performance and by optimizing processes and production. The Industrial Internet offering consists of advanced monitoring and prediction applications which are combined with service agreements where case company experts offer remote services for their customers. The current way of performing remote services in the case company is to utilize cloud-based advanced monitoring and prediction applications for reporting and on-demand

support. The reporting intervals are agreed case by case with customers. The on-demand support refers to a process where a customer initiates an issue where they need support from the case company. Experts from the case company react to these initiated cases and start solving the issues with real-time data, analytics, and product and process know-how. However, the case company has started to develop its offering towards a more proactive approach where instead of reporting in intervals and based on customer need, the experts in the case company will monitor monitored assets in real-time and contact customers proactively before failures happen to prevent unnecessary shut-downs. Unplanned downtime is costly in the process industry which is why it is critical to notice possible failures as early as possible and do preventive maintenance planning for a planned shutdown. Related to this proactive approach, the case company has started to develop further its predictive monitoring capabilities which means utilizing real-time data and analytics to notice potential issues and findings as early as possible. This can be done by creating alarms with real-time data which transits an expert's focus to the most critical assets. The case company has started to develop an edge solution to broaden up their offering to support customers who do not allow cloud connection. The edge solution can not prevent performing remote services which is why the edge solution needs to have the required capabilities for performing remote services. This is the focus of this case study.

Before starting the research, background information about the current situation of the edge solution was gathered. There were sixteen research participants chosen with convenience sampling and additionally, two additional participants were proposed during interviews who did not fill pre-questionnaire but were interviewed. All participants are presented in an anonymized way in Table 2.

Table 2. *Research participant information*

Participant ID	Industry business line
P1	Technologies
P2	Technologies
P3	Automation
P4	Automation
P5	Technologies
P6	Technologies

P7	Technologies
P8	Technologies
P9	Automation
P10	Service
P11	Service
P12	Technologies
P13	Technologies
P14	Service
P15	Service
P16	Service
P17	Automation
P18	Service

The research was conducted in two parts. First, a pre-questionnaire was sent to the participants to determine the maturity level of the remote services and to get a background understanding for the interviewer about the current situation and scope of the remote services. The pre-questionnaire was short, and the questions were easy to answer so that all participants would answer the questionnaire. The questionnaire was answered by 15 out of 16 participants. The questionnaire had ten questions, one background question to determine the business unit where the participant belongs, and one question determining whether the participant agrees that the answers are utilized in this thesis. All participants agreed that their answers can be anonymously utilized in this thesis. The ten questions, their type, and the reasoning why they were chosen are presented in Table 3.

Table 3. *Questions in pre-questionnaire*

<b>Question number</b>	<b>Question</b>	<b>Question type</b>	<b>Significance for research</b>

1	What kind of remote services are you offering?	Multichoice question	This question helps to answer RQ1.
2	Who is conducting the service? e.g. first technology, then internal expert	Multichoice question	This question helps to answer RQ1.
3	What kind of tools are currently critical for conducting these remote services?	Multichoice question	This helps to answer RQ2.
4	Which questions does your service answer?	Multichoice question	This question is used to determine the maturity level of remote services at the moment.
5	What is/are the focus in your service?	Multichoice question	This question is used to determine the maturity level of remote services at the moment.
6	What is the time perspective in your service?	Multichoice question	This question is used to determine the maturity level of remote services at the moment.
7	Which techniques/models are used in your service?	Multichoice question	This question helps to answer RQ2 and this question is used to determine the maturity level of remote

			services at the moment.
8	How does your service support decision-making?	Multichoice question	This question is used to determine the maturity level of remote services at the moment.
9	In some cases, the customer is unwilling to transfer data to the cloud. Do you see that edge solution could boost our chances of winning potential sales cases?	Single choice question	This question tries to determine whether H1 is true.
10	Do you already have a plan for how you would like to develop your remote services in the future? What is the aim? Are new capabilities or data needed to reach these goals?	Open question	This question helps to answer RQ3.

The second part of the research was the interviews. There were twelve interviews, multiple participants were involved in some interviews. The interviews were semi-structured, they lasted one hour, and the interview structure was as follows:

- Introduction to the thesis (5 minutes)
  - o Introductions
  - o Thesis topic
- Going through interesting highlights or unclear parts from the pre-questionnaire (10 minutes)

- Presenting the current situation of the edge solution (10 minutes)
- Discussion about selected themes (15 minutes)
  - o Overall thoughts and questions about the edge solution
  - o Benefits of edge solution
  - o Weaknesses of edge solution
  - o Opportunities for an edge solution
  - o Threats of edge solution
  - o Is there some capabilities in the cloud that would be lost with a single edge solution
  - o Thoughts on hybrid, cloud + edge, model
- Filling questionnaire together (20 minutes)

The discussion part of the interview aimed to answer RQ4 and RQ5 and the questionnaire in the end aimed to gather requirements that were base for answering to RQ7. The questionnaire that was filled together during the interview contained one likert scale question “How important do you see these capabilities from your remote service point of view? (Now-2years)” and seventeen capabilities that were rated on a three-level scale of low, moderate, and high. The questionnaire is in Appendix A. The questionnaire was filled out, and the capabilities were discussed to understand the reasoning behind the rating decision since it is not always straightforward.

In the twelve interviews, few additional names were mentioned outside the scope who had further information or experience about edge solution that were interviewed additionally without filling the two questionnaires to broaden up understanding on the topic. These interviews were included to this research.

## 3. PROCESS INDUSTRY AND SERVICIZATION

This chapter aims to introduce the context, process industry, with its characteristics, challenges, and its prevalent trends. Even though the technology to provide services in the process industry is continuously developing, the characteristics and challenges remain.

### 3.1 Process Industry and its characteristics

The goal in the process industry is to ensure efficient, safe, and resource-saving production with required product quality (Schmidt et al. 2018). The process manufacturing industry differentiates from conventional industries by having cyclic material flows, diversity in the end products' qualities, and non-linear production processes that are not fully controllable (Krumeich et al. 2016). A process automation systems are autonomous, and they are controlled and managed with little or no human effort. The process automation consists of data acquisition for data collection, remote monitoring which indicates system status and alerts from abnormal situations, and controls that operate and control the system. (Xu et al. 2018) The process industry has two types of downtime, planned and unplanned. The unplanned downtime needs to be reduced since it can create costs up to tens of billions of dollars a year. The unplanned downtime can be a consequence of equipment failures. (Azeved 2022) However, even smaller deviations in equipment performance can create harm for the operations and additional costs (Ren et al. 2022). This is why continuous monitoring and optimization of business processes is a must in the process industry to stay competitive. (Krumeich et al. 2016) The common concepts that are monitored are the status, performance, and health of facilities, production lines, and equipment. The continuous monitoring enables data-driven decision-making based on insights obtained from raw data. (Azeved 2022) Additionally, since equipment failure creates costly unplanned downtime when it happens, it is important to make forecasts and do predictions (Angelopoulos et al. 2019) However, it has been difficult to make accurate forecasts. The forecasts have been done statistically since more complex methods require more detailed data collection which is complex, expensive, and time-consuming. (Krumeich et al. 2016)

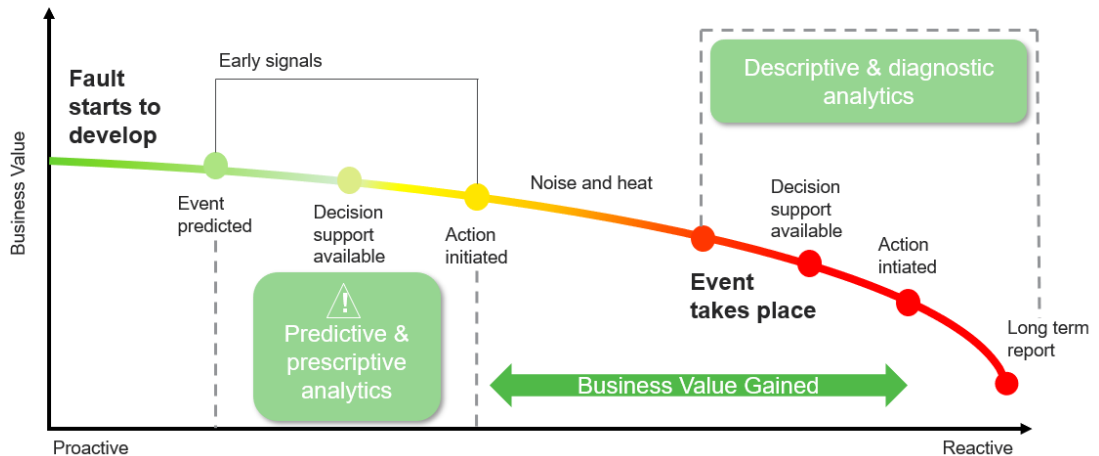
The industry has transformed from traditional automation to the fourth industrial revolution, Industry 4.0, where machines and devices are connected to intelligent systems through the Internet of Things (IoT) meaning identification, detection, communication, and data collection by radio frequency identification and sensors. The main concept of

IoT is to combine physical objects with virtual objects in a network to enable interaction and cooperation towards common goals. (Eigner & Stary 2023) This Industry 4.0 and IoT capabilities have transferred process automation towards fully automated systems where manufacturing processes are regulated based on feedback from sensors and actuators which requires low-latency communication (Jin et al. 2024). However, the goal of Industry 4.0 is not to replace human effort but to combine it with other components in Cyber-physical systems (CPSs) which enable exploiting individual skills and talents successfully (Angelopoulos et al. 2020). CPSs combine hardware, software, computational components, and physical components together to sense and control the state of the real world in real time (Colombo et al. 2017). In the process industry, CPSs refer to the network of raw materials, equipment, humans, processes, and data that are connected through sensing, actuating, computing, and communication technologies (Colombo et al. 2017; Hästbacka et al. 2022). This is made happen with many technological components like sensors, actuators, and networking. Subject matter experts can utilize the collected information and analyze it for decision-making support in real time. The CPS interaction makes the work of the subject matter expert more effective with increased capabilities of the analysis. (Colombo et al. 2017) Since IoT and CPSs have created possibilities to collect data in a cost-neutral way, they provide opportunities for new analytical methods. (Krumeich et al. 2016)

### **3.1.1 Different ways to do analytics and gain business value**

There are four types of analytics in the process industry: descriptive, diagnostics, predictive, and prescriptive analytics (Azeved 2022). These four types of analytics are visualized in Figure 4. Descriptive analytics explains what has happened, diagnostic analytics explains why something happened, predictive analytics predicts when something will happen, and prescriptive analytics gives recommendations for actions to change the predicted outcome (Azeved 2022). Descriptive analytics is conducted mostly as a type-based historical data analysis. This kind of analysis is suitable for reporting and dashboard purposes, and it is the most common analysis type. However, the descriptive analysis can also include real-time data in combination with historical data to explain what has happened and what is the current situation. (Krumeich et al. 2016) Diagnostic analysis seeks to understand why something happened based on historical data. This includes finding root causes and correlations from the data to define a deeper understanding of the relationships between different variables and to define bottlenecks in the process. (Wolniak & Grebski 2023) Predictive analytics strives to predict future conditions like machine breaks. The predictions are based on detecting early signals that are recognized based on rules and patterns that are mined from historical data as shown in

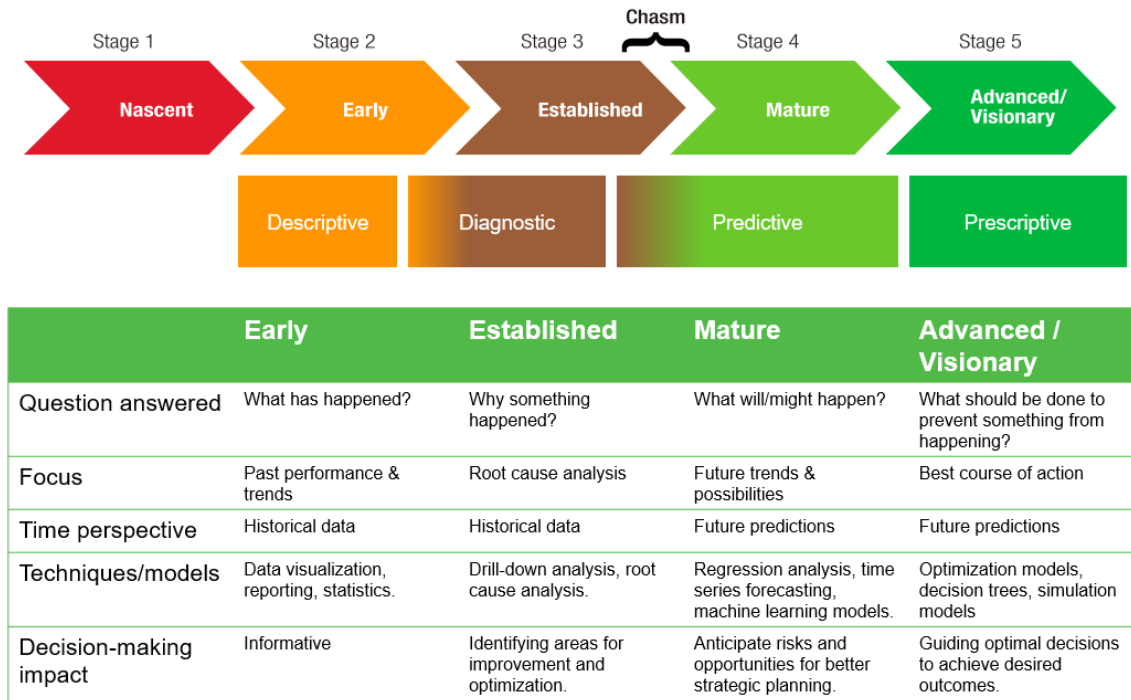
Figure 4. However, predictive analytics doesn't yet guide the corrective actions. For this, prescriptive analytics is required. (Krumeich et al. 2016) The four types of analytics happen simultaneously since they support each other. For example, to be able to make predictions, descriptive and diagnostic analytics are needed to understand what happened and why it happened. (Azeved 2022)



**Figure 4.** Business value in proactive and reactive actions (adapted from Krumeich et al. 2016)

As seen in Figure 4, predictive and prescriptive analytics are proactive and enable proactive maintenance actions like repair, overhaul, and replacement to prevent machine failures from happening. Timely detection and handling of complex events is crucial in means of losing or gaining business value. The earlier the complex event is detected or forecasted, and actions take place, the less business potential is lost as illustrated in Figure 4. Descriptive and diagnostic analytics are good for long term reporting purposes but in these the business value is already lost if they are used alone. (Krumeich et al. 2016)

Based on these four types of analytics, the analytics maturity level of an organization can be determined. These four types of analytics do not however tell the overall analytics level of an organization which includes also resources, data management, governance, and organizational components. (Halper 2020) In this research the analytics maturity level is not the main focus but instead it gives a base understanding where the case company is in analytical capabilities which is why the other aspects than the analytics capabilities are left out from the research. With this decision, an analytics maturity model is created and presented in Figure 5.



**Figure 5.** Analytics maturity model (adapted from Halper 2020; Wloniak & Grebski 2023)

A nascent level is a pre-stage which means that analytics are not utilized yet but interest has begun to rise towards data (Halper 2020). In the early stage, the reports and dashboards are the main tools to handle data which means that descriptive and some diagnostic analytics are in place (Halper 2020; Wloniak & Grebski 2023). In the established stage, reports and dashboards are still the main tool but there is also self-service tools and data scientists who conduct the more advanced analytics. The interest in predictive analytics has started to rise at this point. (Halper 2020) In the established stage the self-service capabilities enable diagnostic analytics and the data scientists might already perform some predictive analytics (Halper 2020; Wloniak & Grebski 2023). Between established and mature stages is a chasm which means that the organization is trying to be more mature but there are some challenges that need to be overcome before getting there like obtaining the right skills, resolving political issues, and choosing the right vision. When the organization overcomes these challenges, they reach the mature level where the analytics is supporting the organization and there is a center of excellence serving different parts of the organization. The analytics might be integrated into business processes and automated and some new technologies like machine learning are in use. However, all collected data might not be fully utilized. (Halper 2020) At this stage, since new technologies are in use, predictive analytics have gained ground in the organization (Wloniak & Grebski 2023). The final stage is advanced/visionary stage where the mindset is to continuously develop the analytics. The connection between existing assets and

new data is well understood and new technological innovations are all the time explored. (Halper 2020) At this stage, the most advanced analytics are in place which means that the analytics level can be already at prescriptive level (Wolniak & Grebski 2023).

### **3.2 Challenges in Process Industry**

The collected data, with IoT capabilities, is currently not fully utilized in terms of business process analytics. (Krumeich et al. 2016) One reason for this is poor data quality. Based on Wang et al. (2020) only 49 percent of collected data is valid and reliable which also makes the analysis results untrustworthy. Additionally, the computing resources are configured and deployed nowadays still in manual and static ways which makes the automation systems inflexible, and deploying new applications or updating the old ones requires a lot of resources and effort (Suter et al. 2019).

In normal operations, automation takes over the process control, and the operator is left with the monitoring task. However, there is little support when a problem situation or starting or shutting down the plant occurs. Sometimes the needed knowledge can be known only by individuals and not be found in the system when needed by an inexperienced operator. Even though operators are skilled, it can be challenging to combine their expertise in an appropriate way. (Schmidt et al. 2018) Hence the importance of utilizing both internal and external information will remain since all available information is never in one place even though the information sharing would be supported and monetized (Isaksson et al. 2018). Decision-making for production is time-critical which is why alarms are commonly utilized in the process industry. However, the alarms can occur simultaneously which can create information overload and make it difficult to identify the most critical ones. (Schmidt et al. 2018) This is another issue where the support would be needed.

There are also only a few solutions for long-term monitoring and optimization. (Schmidt et al. 2018) Also, the automation of fault detection, predictions, and prevention is limited (Angelopoulos et al. 2020).

### **3.3 Services in process industry**

Long-term service agreements (LTSA) are used in the process industry context to create agreements between suppliers and customers about the maintenance of capital-intensive equipment and they can be agreed for multiple years. The idea is to transfer the risk of maintenance from the customer to the supplier. (Bollapragada et al. 2007) From the service provider's viewpoint, the services can increase product sales, increase customer

loyalty, and give better insights about customer needs which can help to improve the offering (Eigner & Stary 2023). Still, the key is to find solutions that create additional value and decrease the customer pain points from which a few major ones were described in chapter 3.2. The service transformation where the product-oriented and service-oriented offerings are co-existing is called servitization. (Kanninen et al. 2017) Previously the most added value has been related to the physical product itself but now the added value is created in combination with the physical product and its related service offering. Since the process industry is still strongly product-focused, the services aim to supplement the product by improving the functionality, performance, and life cycle of the product. (Eigner & Stary 2023) The basis for the servitization is to understand the customer's needs, business, and processes. Previously the customer has been contacted earliest at the sales phase but to achieve a broad understanding, iterative customer collaboration is needed. Servitization also requires an organizational change in mindset from a fully product-oriented view towards a service-oriented view. (Kanninen et al. 2017) Jussila et al. (2017) suggest that there are five service dimensions: pre-sales, product support, product life cycle, R&D, and operational services. Pre-sales services focus on attracting new customers and consist of customer seminars and product demonstrations. Product support services are related to the warranty period and consist of technical user training, consulting, and support services. The product life cycle services consist of installation, repair, and maintenance work which can be done during the whole life cycle of the sold product. The R&D services focus on product development and include for example research and feasibility studies. The operational service dimension focuses on the customer processes and how the product is operated. The services related to the operational dimension are often more radical service innovations. (Jussila et al. 2017)

Service innovation means developing a new service. The radical service innovations are totally new services for the market like predictive maintenance which is part of condition-based maintenance which utilizes real-time data to make maintenance decisions. (Huang et al. 2021; Raddats et al. 2022) In the process industry, the services can be divided into traditional services like maintenance work and to digital services like the previously mentioned predictive maintenance (Gebauer et al. 2021). Predictive maintenance enables decisions about component replacement or repair before the actual failure has happened that would affect the whole production line. Predictive maintenance is based on real-time analysis of the manufacturing equipment and its condition, like overall equipment efficiency (OEE) and early warnings of anomalies, with IoT sensor data that moves the service logic towards a more proactive approach and answers the need for long-term monitoring and optimization in the process industry. (Teoh et al. 2023) Many

services are already recognized in the market like remote monitoring and performance advisory services. Remote monitoring is seen as a critical service nowadays and it also supports the other service offering. Remote monitoring provides customers the possibility to access a high level of real-time support and gain support for failure detection. (Raddats et al. 2022) The lack of this kind of support was identified in Chapter 3.2. Remote monitoring can for example consist of failure detection, product performance monitoring, process performance monitoring, and product condition monitoring. Performance advisory service on the other hand is about consulting the customer for example with performance, efficiency, optimized machine performance, and better product usage-related topics. (Raddats et al. 2022)

## **4. DIGITAL SERVICIZATION IN PROCESS INDUSTRY**

This chapter provides a better description of the digital service transformation existing in the process industry and aims to answer the first research question from the literature point of view.

### **4.1 Digital servitization and smart PSS**

In the consumer market, digitalization has re-formed our everyday lives during the last ten to fifteen years with smartphone apps and the internet. However, in the business-to-business market, like the process industry, the change has not been fully reached even though it is seen that digitalization started already in the 1970s when the first distributed control systems (DCS) were introduced. In the future digitalization will be affecting every aspect of how the process or manufacturing industries conduct their business. The term digitalization refers to the utilization of new types of data, communication infrastructure, and computing power aiming to create new possibilities to create value. Digitalization has gained ground in the process industry since the high-level business objectives, productivity, resource efficiency, and responsiveness, have not changed and the market is demanding a higher flexibility and agility. Customer demands are increasingly more individualistic which requires more flexibility in the offered solutions. Also, the availability and costs of electricity are fluctuating, and the utilization of renewable energy sources has grown in the market. Thirdly, the cost pressure is continuously increasing. To answer these needs, more automatic control and optimization are needed. (Isaksson et al. 2018)

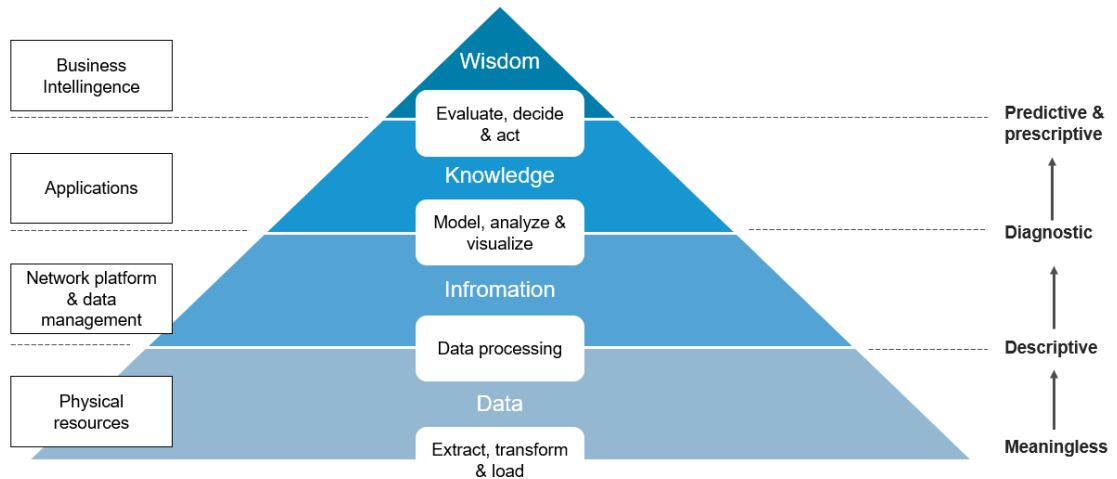
The aim of digitalization in process industry is to gain economic benefits for the industry. The challenges in the process industry can be summarized as two main challenges that are how to ensure 100% availability of the plant and how to engineer systems that are more complex than today with less effort. The 100% availability means the desired situation where there would be only planned shutdowns. This requires a balance between maintenance costs and risk. Availability can also be focused on productivity and resource and energy efficiency. Condition monitoring and predictive maintenance are hot topics with this problem. The latter problem is related to the need to continuously improve the existing systems and at the same time generate new systems, solutions, and products. Both of these challenges are still relevant nowadays and they have been approached with digital capabilities. (Isaksson et al. 2018)

Digital servitization combines digitalization and servitization trends. The digital market opportunities are growing but they have rarely resulted in beneficial outcomes. On the other hand, even though companies are increasing their service offering, the market opportunities are hard due to the high installed base and product life cycle costs. Since these two trends cannot meet the customer needs themselves, these have been combined to create revenue-increasing and cost-decreasing offerings. (Gebauer et al. 2021) Digital technologies create the possibility to offer services over the entire product life cycle. This creates the service provider's competitive advantage, and the customer can benefit from increased flexibility and efficiency of work resulting in time, resource, and cost savings. (Eigner & Stary 2023)

The resulting concept from digital servitization is a smart Product-Service-System (smart PSS) which combines the traditional offering with its digital supplements (Rapaccini & Adrodegari 2022). Before the current smart PSS, the IoT-enabled PSS aimed to offer digital services related to maintenance purposes utilizing IoT capabilities (Zheng et al. 2019). Now the industry is moving towards the smart PSS due to the Operator 5.0 concept where the aim is to enable the collaboration between workers and technology like robots or virtual assistants, utilize technology to improve workers' social sustainability, and to increase utilization of monitoring and prediction capabilities. (Verdugo-Cedeño et al. 2023) This is why Smart Connected Products (SCP), big data, and autonomous interaction between physical and virtual objects are important for the smart PSS (Zheng et al. 2019). In line with Operator 5.0, smart PSS creates value through higher availability of assets, greater flexibility and customization, higher productivity, and people empowerment (Rapaccini & Adrodegari 2022). The SCPs are products that are capable of monitoring, controlling, and optimizing. The products are capable of basic monitoring and alarming level operation but also more advanced predictive and prescriptive intelligence. Even autonomous-level operation. (Rapaccini & Adrodegari 2022)

## **4.2 Data-driven services**

The smart PSS changes the business logic of manufacturers to create value through data-driven services enabling interaction between the service provider and the customer (Rapaccini & Adrodegari 2022). Data-driven services utilize data and analytics to provide services that create value for the customer (Schüritz et al. 2019). There can be seen four stages in data-driven services (Rapaccini & Adrodegari 2022). The concept of data-driven services is presented in Figure 6 where the four stages are seen in the pyramid, needed capabilities in each stage at left and services related to each stage at right.



**Figure 6.** *From meaningful data to wisdom (adapted from Ackoff 1989; Rapaccini & Adrodegari 2022; Zheng et al. 2019)*

The base for the data driven services is the data collected from SCPs with physical IoT resources as shown in Figure 6 (Zheng et al. 2019). The first stage in data-driven services is to extract, transform, and load this data from the installed base. However, this data itself is yet meaningful as shown in Figure 6 at right. (Rapaccini & Adrodegari 2022) The data needs to be processed into information using networking and cloud and edge computing data management capabilities that can include for example analytics, simulations, statistics, and machine learning models (Rapaccini & Adrodegari 2022; Zheng et al. 2019). This processed data, information, can already be used to provide descriptive services. The third stage is to model, analyze and visualize the information to create knowledge that can be used in diagnostic services. (Rapaccini & Adrodegari 2022) At the knowledge layer, there need to be applications and services to deliver the knowledge (Zheng et al. 2019) To reach the predictive and prescriptive level of service, the knowledge needs to be evaluated and used in decision-making to determine needed actions and to create wisdom (Rapaccini & Adrodegari 2022). This requires a clear business model, customer relationship management, and other business intelligence capabilities (Zheng et al. 2019) Even though the stages of data-driven services are standard, the implementation can happen in different ways as illustrated in Figure 7. For simplicity, the second and third stages have been combined.

Responsible →	Technology trusted	Supplier (super-service)	Collaborative (high-touch)	Customer (self-service)
Task ↓				
Extract, transform & load	Technology	Technology	Technology	Technology
Process, model, analyse & visualize	Technology	Supplier	Supplier + Customer	Customer
Evaluate, decide & act	Technology	Supplier	Supplier + Customer	Customer

**Figure 7.** Service types based on who is conducting the service (adapted from Rapaccini & Adrodegari 2022)

There are four different data-driven services. Technology-trusted service is fully automated like self-maintenance, self-prediction, and self-configuration (Angelopoulos et al. 2020; Rapaccini & Adrodegari 2022). This saves time and effort in work. In Managed services -type the service work is done fully by the supplier. This requires a skilled specialist to do the analysis and evaluation work. (Rapaccini & Adrodegari 2022) The Managed service model can be used for example by local service personnel to prepare for scheduled maintenance visits to know which equipment needs maintenance actions. Sometimes the visit might not even be needed if everything can be solved remotely. (Grubic 2018) The benefits for the customer are the same as in the technology-trusted service model but the service provider will also benefit from optimized maintenance visits and reduced warranty costs if the local service personnel can identify that the equipment, they have supplied wasn't the real cause of the issue (Grubic 2018; Rapaccini & Adrodegari 2022). Since Operator 5.0 and smart PSS aim to increase human and technology interaction, some tasks like data processing and analysis can be automated. This applies to all four service models. (Verdugo-Cedeño et al. 2023, Zheng et al. 2019) In interactive services, the analysis and evaluation are done in collaboration between supplier and customer. It is decided in collaboration which party takes the required actions. The benefit is to have more broad expertise in decision-making from both sides. The last service model is data-driven self-service where the supplier provides the customer capabilities to make analysis and evaluation themselves. Actions can be taken by the customer or the supplier. The benefit is utilitarian, but the workers on the customer side also gain advances in skills. (Rapaccini & Adrodegari 2022)

### 4.2.1 Challenges in data-driven services

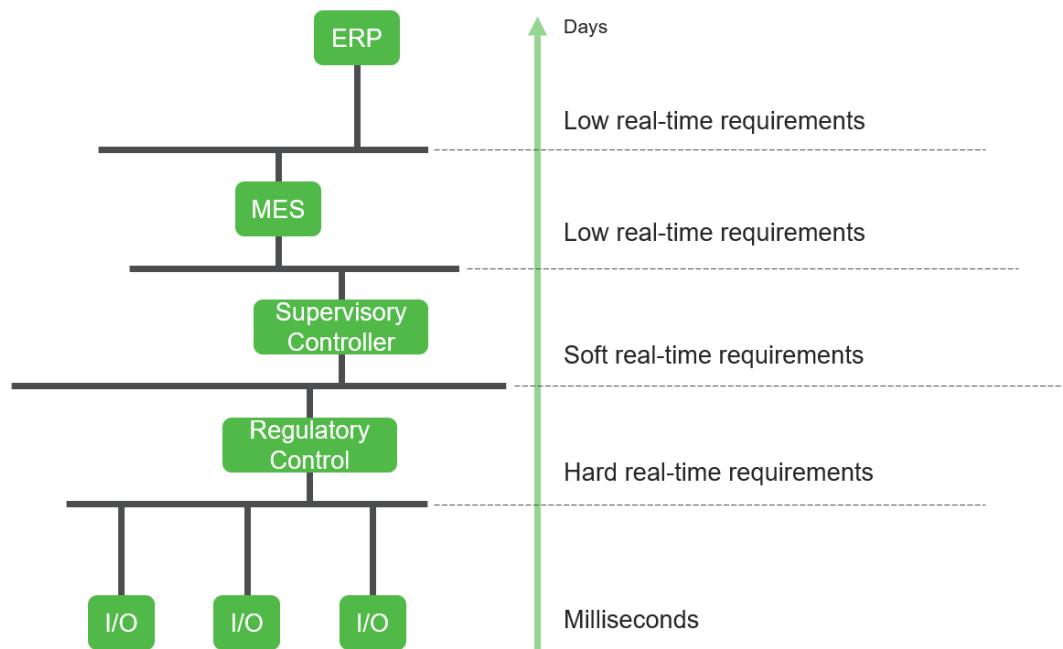
There are still a few challenges when providing data-driven services in the process industry. It needs to be taken into consideration that there are many different customer segments in the process industry that have different needs and they require different service offerings (Kanninen et al. 2017). Similarly, there are many different roles other than just the operator and these different roles have different needs (Schmidt et al. 2018). For example, the operator needs support in time-critical decision-making, but the management needs to be able to analyze a situation over their fleet in the long term (Grubic 2018; Schmidt et al. 2018). Fleet monitoring enables the detection of the most critical assets for maintenance planning, and it also provides the possibility to combine information, for example from problems, between different sites (Grubic 2018). From the technical perspective, it can also be difficult to implement infrastructure for data transfer, storage, processing, and interface with existing plants, brownfield cases, where the infrastructure already exists. Technical problems can also arise when there is a need to incorporate new components into the existing tool landscape. (Schmidt et al. 2018)

### 4.3 Remote services with cloud computing

The traditional automation hierarchy consists of five levels from 0 to 4 as shown in Figure 8. The hierarchy model is associated with manufacturing operations and control systems and other business systems. The level 0 defines the physical production process. Level 1 is for sensing and manipulation of the production processes with sensors and actuators. Level 2 is for monitoring, supervisory control, and automated control of the production process that keeps the process stable and under control. Level 3 is for manufacturing operations management which includes for example reporting, maintenance, quality laboratory tests, movement and storage of materials, and off-line analysis. Level 4 is for business planning and logistics which can include for example planning production capacity, scheduling, preventive maintenance, and utilization of data for maintaining manpower, raw material usage, and spare part usage. (ISA-95.00.01-CDV3:2008, 18) Preventive maintenance detects or prevents degradation of a functional system or component aiming to sustain or extend its useful life (ISO 13372:2012, 1). Time-based maintenance which means scheduled activities based on historical data, and condition-based maintenance mentioned earlier are two sub-categories of preventive maintenance (Huang et al. 2021).

The Industrial Internet of Things (IIoT) has transferred industry to collect and store massive amounts of data from devices (Kaur et al. 2018). In IoT devices are connected to

each other enabling data, information, and resource sharing through the network. Sensors, machine-to-machine collaboration, and other technologies can be used to gather and analyze the data for optimized operations and for providing services. The machine-to-machine interaction is the key to the intelligence of the devices. (Mubeen et al. 2017) Traditionally each layer in the automation hierarchy has been able to communicate only with neighboring layers but IoT enables each layer to communicate and exchange data



**Figure 8.** Automation hierarchy and real-time requirements of each hierarchy level (adapted from Dai et al. 2023; Jin et al. 2024).

with any other layer providing unlimited communication channels.

The data collected from the devices with sensors needs to be processed before it can be provided to the user. This happens traditionally in the cloud (Kong et al. 2022). Hence Sensor-cloud-systems (SCS) have become common which are a combination of wireless sensors and cloud computing (Wang et al. 2020). The SCPs can do simple analysis before cloud computing, but their resources are still limited (Wang et al. 2020, Kong et al. 2022). High power consumption of related devices can also decrease the lifetime of the devices which is why it is not efficient to allocate a lot of tasks for these devices (Kong et al. 2022). Cloud computing has enhanced the CPSs with its capabilities like unlimited storage, high speed of data processing, high computing capabilities, adaptability, virtualization, and multi-tenancy (Armando et al. 2017; Hästbacka et al. 2022; Wang et al. 2020). These characteristics make cloud computing high in performance, scalability, reliability, and cost-efficiency (Hästbacka et al. 2022).

As stated in Chapter 3.1, it is crucial for the process industry that the equipment and machines are operating stably and continuously. Even though SCPs are capable of computing, the capabilities are minor and therefore unsuitable for more advanced computational tasks (Hassan et al. 2018). Fortunately, the cloud computing capabilities have made it possible to compute these more advanced analytical and digital applications like predictive monitoring. Predictive monitoring has been developed to support the monitoring and forecasting capabilities of the equipment and as a result to maximize the equipment uptime (Raddats et al. 2022; Ren et al. 2022). Cloud computing is suitable for these more advanced tasks due to its high computational capacity. The cloud computing logic is that the sensors collect a massive amount of data and send it to the centralized cloud server either through cables or a wireless network. Then the data is processed and analyzed. (Qian et al. 2019)

Cloud computing enables transferring from investments to operational expenses which creates new business models. There has been interest in automation systems that can utilize cloud-based web technologies in the process industry. (Langmann & Mayer 2014). This creates a new hierarchy called Automation Cloud (Jin et al. 2024). In the new automation hierarchy, the critical control and field levels are still formed traditionally at the bottom of the hierarchy. Still, the upper levels, Enterprise resource planning (ERP), manufacturing execution systems (MES), and product lifecycle management systems (PLM), can be migrated to the cloud since they offer non-physical functions and services, they require massive data storage capacity and they have relatively low real-time requirements as presented in Figure 8 (Dai et al. 2023; Givehchi & Jasperneite 2013). In most cases where the traditional automation hierarchy has been accompanied by the cloud, the cloud has been responsible for remote monitoring, increasing business intelligence, and creating short- and long-term reports since it cannot be responsible for low-level procedures that require faster and predictable responses (Mubeen et al. 2017).

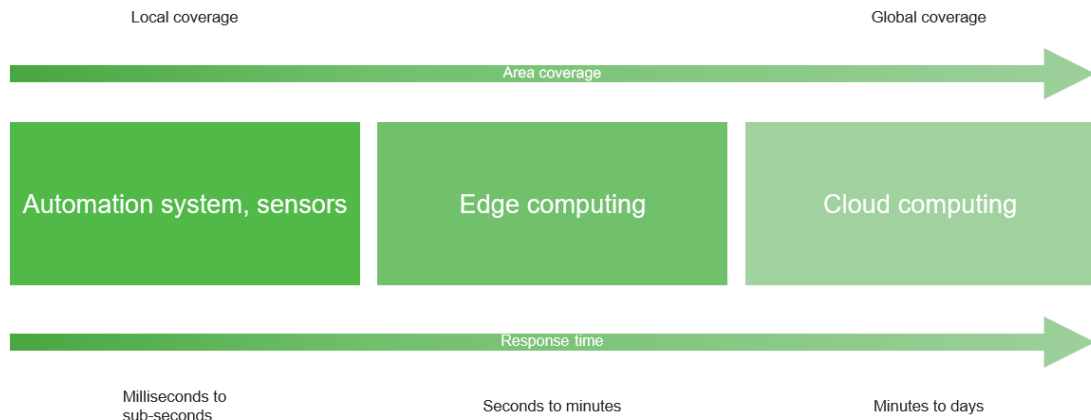
## 5. REMOTE SERVICES WITH EDGE COMPUTING

This chapter aims to explain the edge computing paradigm. This chapter contributes to answer the research questions 3-7.

### 5.1 Edge computing

Cloud computing has been proven to be an efficient way of processing data (Shi et al. 2016). However, when it comes to a massive amount of data to be processed, cloud computing cannot meet the latency requirements (Liu et al. 2019). The response time gets slower when going from the automation system level to the cloud level as presented in Figure 9 where also the coverage level is illustrated in each level (Mubeen et al. 2017). If a user request or a sensed event needs to be responded to within milliseconds, the data must be processed locally. Cloud resources can be used for further processing and long-term analytics. (Suter et al. 2019) The amount of data is continuously increasing in the process industry since IoT applications produce massive amounts of data themselves (Loghin et al. 2019; Shi et al. 2016). Due to this, the speed of processing and especially bandwidth have become bottlenecks for cloud computing which also causes poor user experience (Liu et al. 2019; Shi et al. 2016). Data transfer, storage, and analysis in the cloud also create costs which is why it might not always be feasible to transfer all data to the cloud (Hästbacka et al. 2022). In previously mentioned cases, edge computing can help to reduce latency, network pressure, needed bandwidth, and cloud resources and costs (Shi et al. 2016, Qian et al. 2019). Additionally, edge computing is characterized by proximity, location awareness, geographical distribution, and network context information (Hassan et al. 2018). Since the systems are becoming more complex, it has been observed that cyber-physical systems of systems require novel distributed management, optimization, and control methodologies that can also handle autonomous systems, human interaction, and changing system structures. (Isaksson et al. 2018) Edge computing is a computing paradigm where the computing happens at the network edge instead of centralized cloud servers (Qian et al. 2019). Edge itself refers to any computing and networking resource like servers, base stations, and routers, that are located between the cloud and the physical thing (Hassan 2018; Shi et al. 2016). The key technologies in building edge solution are not static but a basic framework for edge solution, based on popular edge computing platforms, contains Open-Source components, platform deployment components, cross-platform support components, edge

computing enhancement components and edge computing application components (Ning et al. 2020).



**Figure 9.** *area coverage and respnse time in automation systems, edge and cloud (adapted from Mubeen et al. 2017)*

As already highlighted, one of the basic reasons why edge computing is preferred over cloud computing is low latency and the basic idea is to provide better services for the end customer and to increase the network efficiency (Hassan et al. 2018; Taneja & Davy 2017). Applications that require short response time, delay sensitive applications, have become common in process industry which creates new requirements like real-time analysis and real-time decision-making support (Hassan 2018, Loghin et al. 2019). The fault diagnosis of rotating machines for example is highly time-sensitive since these rotating machines, like compressors and motors, are the cornerstone of the processes and they are prone to failures (Qian et al. 2019). Edge computing can answer these requirements since achieving real-time analysis in the cloud is nearly impossible (Hafeez et al. 2021). Another driver for edge is a desire to increase customer authority in data and application management. This enables customers also to operate private data in edge. (Ning et al. 2020) Third driver, related to the authority, is that some industrial sites prefer to keep data on site due to privacy and confidentiality concerns and sometimes the data is simply infeasible to transfer to the cloud for example due to high costs (Suter et al. 2019).

Edge computing still has some shortcomings compared to cloud computing. The most crucial shortcoming is the limited resources for computing which is why tasks that require high computational resources should not be distributed to the edge (Ashouri et al. 2018). Edge devices can have high energy consumption if energy efficiency is not considered and the tasks that require high computational tasks can increase the energy consumption of edge devices (Kaur et al. 2018). Another shortcoming compared to the cloud is that it is not possible to combine data from multiple sources at the edge (Hafeez et al. 2021). It is seen important to start integrating production planning with procurement of energy and raw material and with maintenance planning to ensure that both utilization

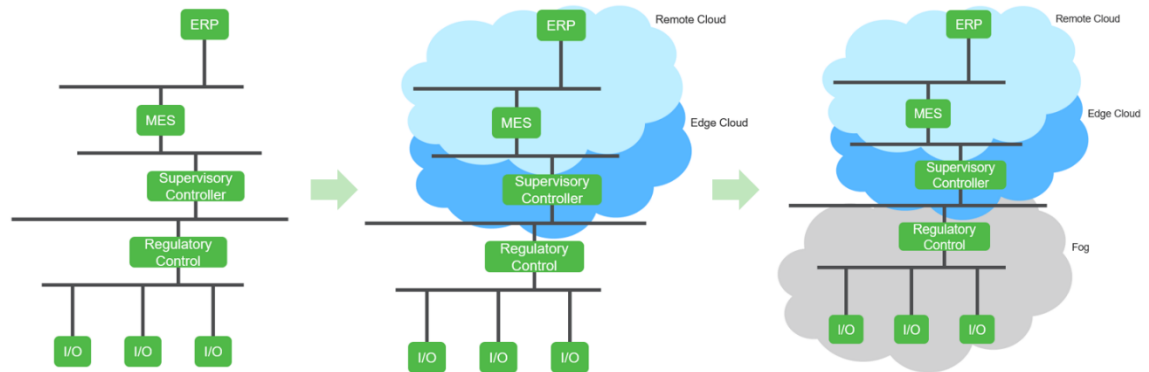
and availability are considered, where it is critical to be able to combine data from multiple data sources (Isaksson et al. 2018). As mentioned in Chapter 3.2.1 the analysis of the whole fleet is required to get an overall view of the situation. Edge computing is characterized as secure and private, but it is still new computing paradigm and therefore it can bring new and unknown security and privacy problems (Kong et al. 2022). One data security problem that need to be solved at the edge is how to ensure data security when multiple edge devices cooperate with each other and how to not expose users' location information (Ning et al. 2020). Finally, from the technical perspective, the cloud infrastructure is transparent to the user but on the edge, the platforms are heterogeneous which makes it difficult for the programmers to write code that can be deployed in different edge computing platforms. Additionally, the collected data comes in different formats from different equipment and there is not yet a standardized mechanism for the formatting in edge which also makes the work of the programmers more difficult. (Shi et al. 2016)

Regardless of these shortcomings, utilization of edge computing will grow especially in smart homes, smart factories, and smart city contexts. Additionally, there has been discussion about intelligent edge referring to the utilization of artificial intelligence technologies at the edge for processing data. (Ning et al. 2020)

## **5.2 Edge cloud hybrid model**

Since cloud and edge have many shortcomings alone, a hybrid model has been proposed to utilize the benefits of both. The idea is that the edge can take partially tasks from the cloud server when it is applicable for example to decrease the burden of the cloud or to answer the low response time requirements. (Kong et al. 2022) Edge computing can be seen as an extension of the cloud, but it cannot replace the cloud completely since some applications still require cloud computing (El-Savyed et al. 2017; Hästbacka et al. 2022). The edge cloud hybrid model benefits from the large computing capabilities of the cloud but is still able to answer the requirements for low latency (Hassan et al. 2018). Monitoring and control of equipment are examples of tasks in the process industry that usually require short response times. (Hästbacka et al. 2022). As a result, the edge cloud hybrid model provides a trade-off between latency, bandwidth, availability, reliability, and scalability and provides different computing options (Tusa et al. 2024). Additionally, security can be increased, and power consumption can be decreased (Kong et al. 2022).

In addition to the previously mentioned Automation Cloud, it has been envisioned (Jin et al. 2024) that the supervisory control could be also deployed locally in edge as part of Cloud Automation, presented in Figure 10 in the middle to answer the lower latency and security requirements.



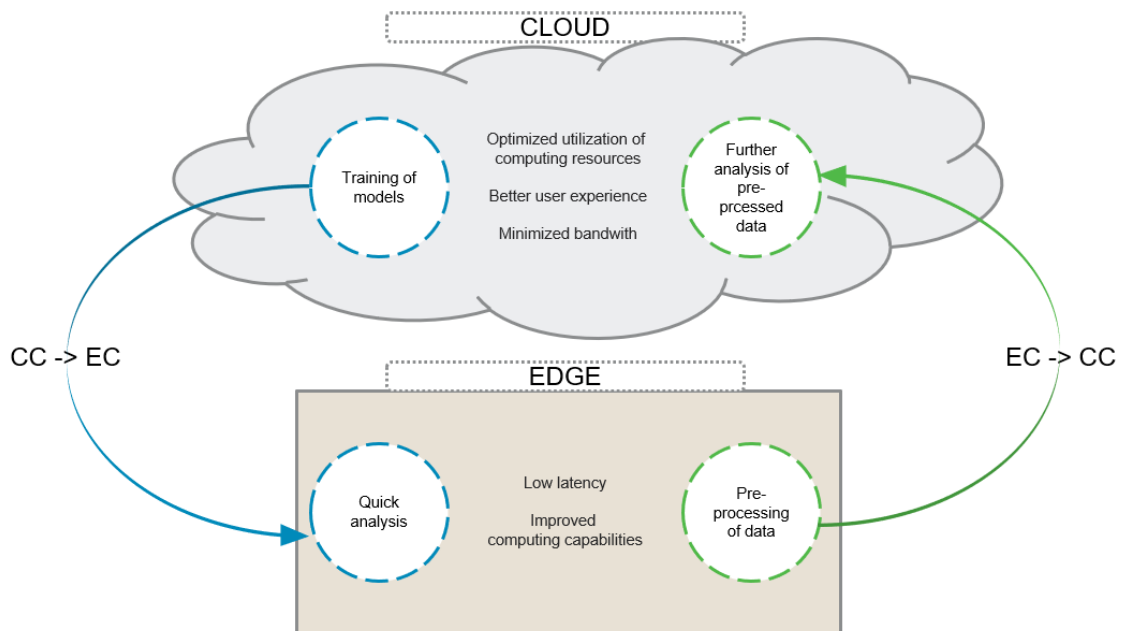
**Figure 10.** Edge and cloud potential in automation hierarchy (adapted from Jin et al. 2024)

It has been envisioned that the lower-level machine control and physical machine level could be deployed in fog, referring to Cloud-Fog Automation which is illustrated in Figure 10 at right (Jin et al. 2024). Fog and Edge are often used as synonyms in literature, but Jin et al. (2024) differentiate Fog from Edge so that Edge Cloud is connected to the Cloud, and Fog is separated from the Cloud and can work fully independently. However, the definition is not standardized. El-Sayed et al. (2017) define fog computing as a combination of IoT, edge, and cloud. Based on Ning et al. (2020) edge computing and fog computing are similar in that sense that they both bring cloud computing capabilities to the edge of the network. However, they define that edge computing focuses on the things level and fog computing on the infrastructure level and that edge computing happens in edge devices and fog computing happens near edge devices and core networking devices (Ning et al. 2020). In Cloud-Fog Automation based on Jin et al. (2024), the lower-level connections can be changed from wired connections to wireless connections which brings new opportunities for IoT use case and deployment. However, it has been hard to get the same reliability, latency, security, and functionality as with wired connections (Jin et al. 2024).

### 5.2.1 Different layers of the hybrid model

In the smart PSS, there can be seen three layers: physical, cyber, and system layers. The physical layer is responsible for sensing, computing, and communication and consists of SCPs and related hardware like sensors and actuators. The cyber layer is re-

responsible for data preprocessing, storage, and analysis and utilizes edge and cloud computing. Finally, the results are further analyzed, visualized, and reported in the system layer. (Liu et al. 2019) In the hybrid model architecture, the cyber layer is separated into an edge layer and a cloud layer. The hybrid model consists of thing, edge, and cloud layers. The thing layer is responsible for sensing, actuating, and communicating consistent with the physical layer of smart PSS. However, edge and cloud can have different roles and responsibilities depending on the situation. The edge layer can process and analyze data and interact with both, thing, and cloud layers. The edge layer can even host simple IoT applications. Cloud can instead host multiple IoT applications requiring high computational capacity. (Ashouri et al. 2018) Most of the collected data is temporary or even useless which is why it should be pre-processed before transferring it to the cloud (Hafeez et al. 2021; Ning et al. 2020). After pre-processing, only the required data should be sent to the cloud with less bandwidth and minimized costs for further analysis which also improves the user experience in IoT applications since the response time gets better (Hafeez et al. 2021; Liu et al. 2019; Ning et al. 2020). This model, EC->CC, is illustrated in Figure 11 at left. Additionally, there is another edge-cloud-interaction model, CC->EC that is presented in Figure 11. CC->EC integration answers the latency issue. Since edge doesn't have enough resources for training advanced analytical models, it is beneficial to utilize the cloud computing resources by training the models in the cloud and then transferring the trained model parameters to edge for local quick analysis. (Hafeez et al. 2021; Liu et al. 2021)



**Figure 11.** Edge-cloud-integration model (adapted from Liu et al. 2021)

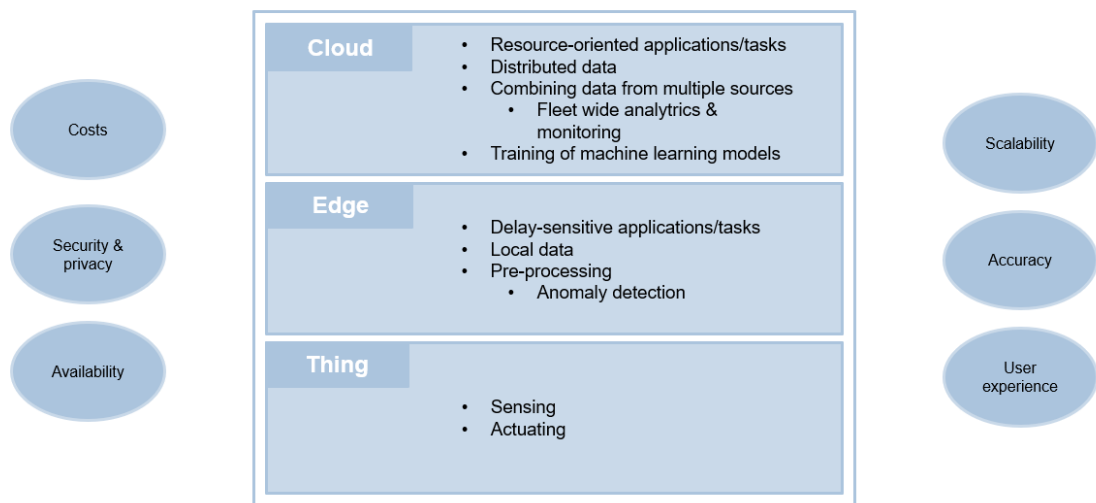
Predictive maintenance is a common application in the process industry, as mentioned previously. However, Predictive maintenance requires a massive amount of data to be

processed and analyzed with machine learning models. Traditionally the data have been sent to the cloud for processing and analysis, but the data transfer and processing has been proved to be costly. Edge on the other hand doesn't have enough resources for this kind of computing. (Hafeez et al. 2021) In the case of Predictive maintenance, the EC->CC and the CC->EC logic can be used. First, the data can be collected to the edge for pre-processing (Bellavista et al. 2020). The pre-processing can include for example data cleaning or enlargement (Ren et al. 2022; Wang et al. 2020). Sometimes there might be a lack of historical data on the predicted condition or the condition might be changing. This problem can be resolved with data enlargement where the gaps in the data are filled. Sometimes, there might be multiple faults that are happening simultaneously and have relevancies. This can be utilized in data enlargement by letting the model adopt samples from other conditions. (Ren et al. 2022) After pre-processing, the processed data can be used in a local edge version of the machine learning model that has been previously trained in the cloud, CC->EC. This enables real-time decision-making. From the edge, the processed data can be forwarded to the cloud where the machine learning model can be further trained, EC->CC. The model could be trained in edge, but it would consume a lot of computing resources. (Bellavista et al. 2020)

### **5.2.2 Placement decision in hybrid model**

It has been proved by Taneja & Davy (2017) that the edge cloud hybrid model can decrease latency and network usage but the placement decision, where to place each component in the architecture, affects these. Wrong decisions in architecture can even increase costs and decrease performance. It can be difficult to decide where to compute each component or task, but the solution should maximize the performance and minimize costs. Since the thing layer in the hybrid model is capable of sensing and actuating, these tasks should be happening in this layer. Ashouri et al. (2018) and Kaur et al. (2018) propose that in the case of other tasks, the tasks should be categorized into delay-sensitive and resource-oriented tasks. Delay-sensitivity means requirements for low latency and resource-oriented nature means tasks that require high computational resources. The delay-sensitive tasks should take place in the edge and the resource-oriented tasks in the cloud. Kaur et al. (2018) and Zheng et al. (2019) agree that small-scale, short-term, and local data should be processed in the edge and large-scale, long-term, and distributed data can be stored and analyzed in the cloud. It is important to acknowledge that the network devices can get overloaded if the interaction between the edge and cloud is highly active which increases the latency (Kaur et al. 2018). Ashouri et al. (2018) propose that tasks should be computed in edge if they don't require any interaction with the cloud or if the interaction can be done with optimized performance and costs.

There are still a few components that should always be computed in either edge or cloud. Fleet-wide analytics and monitoring should always happen in the cloud since it enables data collection and comparison from multiple locations (Hästbacka et al. 2022). Also, more advanced models like machine learning should always be located in the cloud even though they can have a local version in edge (Hafeez et al. 2021; Hästbacka et al. 2022). On the other hand, pre-processing, like data cleaning, should always happen in edge (Hästbacka et al. 2022; Wang et al. 2020). Since the data is heterogeneous and noisy, traditional data cleaning methods are not suitable. Anomaly detection is often used with this kind of data. Anomaly detection detects abnormal data points with its statistical algorithm which can be based on distance or density. (Wang et al. 2020). Anomaly detection is part of condition-based maintenance, and it monitors equipment in real-time. Anomaly detection should be in edge to decrease the energy consumption of the sensor nodes and to increase the bandwidth of the data transfer to the cloud. (Hästbacka et al. 2022; Wang et al. 2020). Additionally, since edge can execute Anomaly detection near the automation system, it can achieve real-time response (Huang et al. 2021).



**Figure 12.** Framework for placement decisions in cloud, edge, and thing levels (adapted from Hästbacka et al. 2022)

As illustrated in Figure 12, there are still many other attributes that affect the placement decision. Accuracy means the difference between expected and actual value. Edge accuracy can be lower than in the cloud with more advanced models since the processing and storage resources are lower. Availability is related to response time meaning how accessible the solution, like an IoT application, is. (Ashouri et al. 2018) Security and privacy have been a big concern when it comes to the cloud but edge on the other hand can bring new security and privacy problems. (Ashouri et al. 2018; Kong et al. 2022) When it comes to costs, there are different costs, implementation costs, and operational

costs (Ashouri et al. 2018). The operational costs can be divided into energy usage, network characteristics, application characteristics, and monitoring-related costs (Francis 2018). Related to implementation costs, edge solution has substantial upfront investment expenses (Hassan et al. 2018). Scalability includes, for example, the ability to adapt to workload changes, and elasticity. Poor elasticity can affect costs, performance, accuracy, availability, and latency. (Ashouri et al. 2018) User experience is affected by response time (Liu et al. 2019). The placement decision needs to be done case by case.

### 5.3 Cloud, edge, or both?

Single-edge, single-cloud, and hybrid model comparisons are summarized in Figure 13. As can be seen, the benefits of the single-edge solution are low latency, reduced need for bandwidth, and high reliability. Energy consumption is usually low with edge devices, but it can get high with resource-oriented tasks. On the other hand, limited computing resources and storage are shortages of the single-edge solution, as well as the limitation of combining data from multiple sources. The single-cloud solution has high computing resources, low energy consumption, unlimited storage, high reliability, and the possibility to combine data from multiple sources as benefits. High latency and the need for high bandwidth are shortages of the single-cloud solution. Finally, the hybrid model has all the same benefits as the single-edge and single-cloud combined but as a risk, wrong placement decisions can increase latency and energy consumption more than in single-edge or single-cloud solutions.

	Edge	Cloud	Hybrid
Latency	Low	High	Optimized (wrong placement decisions can increase latency)
Bandwidth	Reduces need for bandwidth	Bottleneck with massive amount of data	Optimized
Computing resources	Limited	High	High
Energy consumption	Low (High with resource-oriented tasks)	Low	Optimized (wrong placement decisions can increase energy consumption)
Combining data from multiple sources	Not possible	Possible	Possible
Storage	Limited	Unlimited	Unlimited
Reliability	High	High	High
Distribution	Geographical	Centralized	Geographical

■ Shortage    ■ Risk

**Figure 13.** Comparison between single edge, single cloud, and a hybrid model

It is not always straightforward to compare the differences in single-edge, single-cloud, and hybrid model. Security and privacy, costs, scalability, and performance are examples of features that are different in different situations. Security and privacy have been

a major concern in the case of single-cloud solution. Matter in fact, this has been the prime reason why organizations have been reluctant to transfer their data to the cloud. (Ometov et al. 2022) The two barriers to cloud adoption from privacy, security, and trust perspectives are fear of leakage of sensitive data to competitors, and loss of privacy. There are many aspects of the privacy concern: lack of user control, fear of unauthorized secondary usage, data proliferation, and data flow to third-party users. When the SaaS cloud environment is used, the service provider will gain control over the data storage resulting in a lack of control for the customer which creates fear of data theft, misuse of the data, and unauthorized resale of the data. One concern is whether the data is deleted from all places if the customer requests since it can be hard to guarantee that all copies or backups of the data are deleted. These concerns can be decreased with clear and explicit statements in the agreement about the data protection responsibility of the service provider, liability for failure, geographical regions where the data is located, compliance of applicable data protection and privacy laws, ownership of the data, what will happen when the agreement is ending or if the data is lost and defining policies for data retention and deletion. (Pearson & Benameur 2010) However, the edge also has security and privacy risks related to data transfer, vulnerabilities of edge nodes, and access control. There are similarities and differences in these known risks. (Ometov et al. 2022) However, edge computing can bring new and unknown security and privacy risks (Kong et al. 2022). As a similarity, in both computing paradigms, there is a risk that the wrong person can get access to private information and data. Also, the data transfer is vulnerable to attacks in both computing paradigms. In edge, one security and privacy risk is related to the vulnerability of the edge nodes since an attacker can push harmful data to the devices. In the cloud, this risk is not present but there is a risk of losing data due to high traffic and usage of cloud. (Ometov et al. 2022) Since the hybrid model is a combination of the single-edge and single-cloud solution, it has all the above-mentioned security and privacy risks (Pan et al. 2018).

As mentioned earlier, costs can be divided into implementation and operational costs which can be divided further into energy usage, network characteristics, application characteristics, and monitoring-related operational costs (Ashouri et al. 2018; Francis 2018).

Energy consumption is usually low with edge computing. However, it can get high if the energy efficiency is not considered and the tasks that require high computational tasks are located at the edge (Kaur et al. 2018; Shi et al. 2016). Data transfer to the cloud creates high energy consumption. In the hybrid model the energy consumption can either be minimized with correct placement decisions but on the other hand increased with wrong one. (Ashouri et al. 2018)

From the data transfer perspective, cloud computing creates larger data transfer costs since the data centers are located far away from the data source (Hästbacka et al. 2022). When it comes to the edge, computing happens as close to the data source as possible which can reduce the data transfer costs (Shi et al. 2016). In the hybrid case, the data transfer costs can be reduced since the data can be already pre-processed at the edge and fewer data can be sent to the cloud. However, it is important to acknowledge that network devices can get overloaded if the interaction between the edge and the cloud is highly active (Kaur et al. 2018). The costs of edge computing arise from implementation and maintenance and in cloud computing from data transfer, storage, and analysis (Hassan et al. 2018; Hästbacka et al. 2022).

Cloud is characterized as scalable but there can be seen different kind of scalability (Dastjerdi & Duyya 2016; Hästbacka et al. 2022; Kong et al. 2022). There is computational scalability and network scalability. The cloud is scalable in computation and storage. Since the cloud is centralized, sending a lot of data to cloud can saturate network bandwidth and as a result reduce network scalability (Dastjerdi & Buyya 2016; Kong et al. 2022). Edge on the other hand is geographically distributed which is why the network scalability in edge is high (Kong et al. 2022). However, the edge is not scalable in sense of computation and storage due to its limited resources (Dastjerdi & Buyya 2016). Since the hybrid model can benefit from both edge and cloud characteristics, the hybrid model is scalable in computation and storage and high in network scalability (Kong et al. 2022).

Loghin et al. (2019) made research about performance trade-offs in single-edge, single-cloud, and hybrid model processing. They excluded data collection and post-processing from their research and focused only on the main processing and data transfer between edge and cloud. Also, there was an assumption that the data needs to be always post-processed in the cloud which might not be the case in all situations, especially in an industrial context. As a result, there was no clear winner but the application characteristics like ratio between input data and output data size and bandwidth were affecting the performance the most. When comparing single-cloud and single-edge computing, edge computing is performing better when the bandwidth is low except if the input data size is small. This is because the smaller the input data size is, the less bandwidth is needed in cloud computing. When comparing single-cloud computing and hybrid model, cloud computing is always better in performance except if the input data size is large and the output data size is small. (Loghin et al. 2019)

## 6. RESULTS

This chapter presents the results from pre-questionnaire and interviews. 15 out of 16 participants responded to the pre-questionnaire.

### 6.1 Pre-questionnaire

The aim of the pre-questionnaire was to understand what kind of services are provided and what kind of technical capabilities are required for conducting remote services in the process industry. The first three questions from the pre-questionnaire aimed to answer these questions. There was an existing hypothesis about what kind of remote services there is in the case company from which the first question was formed. The hypothesis was that there are cloud-based IoT applications, reporting services, on-demand support services, and 24/7 services. Based on 15 answerers, the most common service portfolio consists of cloud-based IoT applications (12 answers), reporting services (13 answers), and on-demand support services (13 answers). 24/7 support is not yet widely utilized (3 answers) and there was additionally one other type of service, analysis services (1 answer).

The second question aimed to understand how the services are conducted based on Figure 7. Almost in all services (14/15) an internal expert is performing the service. In more than half of the services (9/15) the service is conducted in collaboration with customers. Only in (3/15) services, the customer is conducting the service in a self-service manner. It was seen by only (6/15) respondents that technology is performing part of the service. In interviews, this decision was reasoned by the fact that the technology can not yet provide the service without human effort even though the technology is collecting the data and pre-processing it.

The third question aimed to deepen the understanding of how remote services are performed and what technical capabilities are required to perform these services. The tools to choose from, in the multiple-choice-question, were based on existing hypotheses on what kind of tools are required. The most utilized tools are DCS screens with remote access (11/15) and the case company's own digital IoT platform (10/15). Other commonly used tools are Excel (8/15) and the case company's customer portal (7/15). In addition to the existing hypothesis, there was an "other"-option where Python and commercial statistical tools were mentioned.

The ninth question challenged the existing hypothesis that the edge solution is required since customers are increasingly unwilling to transfer their data to the cloud. Based on the answers, 5 out of 15 respondents saw that the edge solution could help in sales cases, 4 out of 15 respondents saw that the edge solution could sometimes help in sales cases, 5 out of 15 respondents were unsure and 1 out of 15 respondents saw that the edge solution could not help in sales cases.

The tenth question aimed to determine possibly existing development plans for the remote services and the required capabilities for that. This was an optional open question which ten out of fifteen respondents answered. The answers are presented in Table 4 and the common items that were mentioned are highlighted in blue.

Table 4. *Development plans for remote services in the case company*

Respondent ID	Answer
P1	More systematic and more <b>proactive</b> and predictive approach. Identify possible problems before they escalate
P2	Be more predictive, more <b>proactive</b> .
P3	I think this service needs to be further developed. I hope that the pending development will give more possibilities for improving service profitability.
P4	"Target: 1) use cloud data more effectively than nowadays (increasing the number of persons working on data). 2) alert-based actions (moving away from reactive work and actions with predefined intervals to more <b>proactive</b> ways)  Enablers and needs: new self-service capabilities to allow a "self-sufficient" way of working, without always calling devops for help

P5	Yes, many plans. Improve internal processes, widen knowledge, <a href="#">more advanced analytics</a> ... The aim is to maintain a frontrunner position in remote services while enhancing customer satisfaction and response speed.
P6	Mainly new capabilities are needed. <a href="#">More analytical tools</a> integrated with the platform and also more visualization features. <a href="#">Fleet comparison</a> on different levels mill, machine, process, etc...
P8	We have a long journey ahead of us to widen the analysis support for all of our device types and to develop ticketing and reporting capabilities in the customer portal to be able to run the services more efficiently.
P10	More <a href="#">proactive</a> mathematics is required to provide real-time value for the customer. <a href="#">Data sources need to combine more</a> than currently. <a href="#">More advanced analytics</a> into our analytical applications. Long-time trending. <a href="#">AI-supported</a> automatics leads from our analytical applications. Multi-data source dashboards for customer's operator room.
P14	1. Very basic level should be very cost-effective to keep it online for almost every customer. Examples show that basic-level data analyze does not need to be so sophisticated, those are also needed but that can be something extra then.

	<p>2. In our technology we see great potential to bundle Industrial Internet with other services as we have already done at one customer to <a href="#">combine performance data, laboratory analyses, and inspection</a> at site to make an overall suggestion to the customer. In the future such could be more automatic.</p> <p>3. In long term I expect <a href="#">AI</a> to analyze the data further and find correlations between plants and situations</p>
P15	<p>Combined condition/performance indication, <a href="#">comparable with other mills</a>. Possibly personal/internal alerts for selected process values to notice ASAP if attention is needed. Selected data download to Excel and/or <a href="#">better ways to process numerical data</a> (eg. compare different years) in the cloud. Comparison to reference data available to all visualizations. Better/unified naming /grouping per equipment of tags to be able to find them more easily.</p>

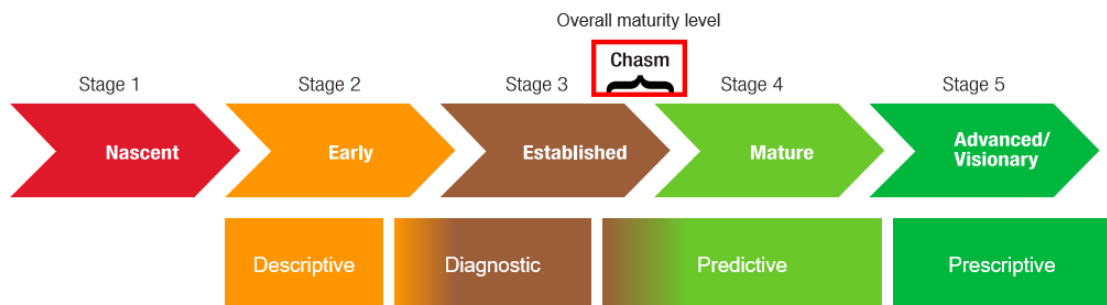
The more detailed answers to these questions can be found from Appendix B. The rest of the questions in the pre-questionnaire aimed to determine the analytical maturity level of the case company as a base understanding of the current situation of the case company and to determine possible future development directions.

### 6.1.1 Maturity level model questions

Based on the fourth question, the services in the case company aim to answer all four levels: what has happened (11/15), why something happened (14/15), what will/might happen (12/15), and what should be done to prevent something from happening (15/15). However, based on the fifth question, the focus of the services is on defining past performance and trends (12/15) and performing root cause analysis (11/15) and less on detecting future trends and possibilities (5/15) and defining the best course of action

(8/15). Also, based on the sixth question, the time perspective in services in the case company is in historical data (15/15) and less in future predictions (6/15).

The techniques/models in the maturity model were separated from each other for the seventh question and there was also "other"-option. The most commonly used techniques/models are reporting (13/15), data visualization (12/15), root cause analysis (11/15), and statistics (8/15). Only a few respondents chose the techniques/models from the highest levels of the maturity model: machine learning (2/15), regression analysis (6/15), forecasting (1/15), decision tree (1/15), and optimization models (1/15). One respondent chose the "other"-option and mentioned one commercial analytical tool. Finally, based on the eighth question, the services in the case company support decision-making by giving information (14/15) and by identifying areas for improvement (12/15) and only partly anticipate risks and/or opportunities (7/15), and guide optimal decisions to achieve desired outcome (7/15). When comparing all these presented results to the maturity model presented in Figure 5, the case company is in a chasm, as presented in Figure 14.



**Figure 14.** *The maturity level of analytics in the case company*

The service is at predictive and prescriptive levels, the focus is on descriptive and diagnostic levels, time perspective is on descriptive and diagnostic levels, techniques that are used are mainly at descriptive and diagnostic levels even though some predictive and prescriptive techniques are in use. Finally, the decision-making support is on descriptive and diagnostic levels. The case company seems to be in transit from a reactive approach to a proactive approach but the processes are not yet standardized and there are some challenges that the case company must overcome to pass the chasm.

## 6.2 Interview results

The main results arise from the "Discussion about selected themes" and from the "Filling questionnaire together" part of the interview. The discussion about selected themes focused on the strengths, weaknesses, opportunities, and threats of a single cloud, single

edge, and DCS system from a remote service perspective. Additionally, the hybrid model was discussed. In the questionnaire, the most important capabilities from a remote service perspective were identified.

### 6.2.1 Single cloud strengths, weaknesses, opportunities, and threats

The single cloud has been the main solution to perform remote services in the case company. The strengths, weaknesses, opportunities and threats of the cloud solutions based on the interviews are presented in Table 5.

Table 5. *Single cloud strengths, weaknesses, opportunities, and threats*

Strengths	Weaknesses	Opportunities	Threats
Easy to access	Sometimes customers do not want to transfer their data to cloud	Fleet monitoring	Losing potential customers
Efficient work	It is not transparent what data is transferred to the cloud	Combining data from multiple sources	
Reference management	Latency	Tasks that require high capacity	
Combining online and offline data	High data processing costs	Training more advanced analytical models	
Possibility to export data			

The main **strength** seen in the single cloud solution based on eight participants was the easy accessibility of the solutions and analytical tools in the cloud which was seen as an enabler for efficient work. P2, P5, and P12 highlighted the importance of collecting reference data to the cloud for reference management purposes. P6, P7, and P13 explained the importance of combining online and offline data, for example when doing troubleshooting, that is possible in the cloud. P18 liked the possibility of exporting data from the cloud since sometimes the data needs to be analyzed in other systems. Overall,

the cloud was seen as a beneficial tool for long-term monitoring and reporting purposes by P17.

In addition to the current strengths, there was seen many development **opportunities** in cloud. The most common, mentioned by eight participants, was fleet monitoring. Another opportunity mentioned by P4 and P10 was a possibility of combining data from multiple sources:

"We don't for example have the maintenance event data in the cloud at the moment. ... So, if we are not even that far that we are able to have the event data and and approach data together in the cloud, we are just at the beginning at the moment, and this should be definitely something to develop further." (P4)

P17 mentioned large capacity needs like training large language models in the future that can be done in the cloud but for example in edge, this would require too much capacity.

Even though there was seen a lot of strengths and opportunities in cloud, there was also many **weaknesses** and **threats**. The main reason for edge development based on P5, P13, P16, and P17 has been the difficult sales situations when the customer has been interested in the Industrial Internet offering but has asked for the local implementation that has not been an option. P13 reasoned that it is not as transparent for the customers what data is transferred to the cloud which might make the customer unwilling to transfer any data to the cloud. Another driver for the local implementation based on P5, P9, and P10 was latency since some IoT applications are delay-sensitive and hence suffer from latency on the cloud side.

"...for some some applications that's going to be amazing because going to the cloud ends up adding like a few minutes and if you've caused a few minutes of delay on that application, it's not as useful." (P9)

However, P2 had a different point of view and highlighted that the case company is not serving the customer in real-time but that the remote services focus on analyzing performance, finding future possibilities, and giving suggestions for process improvements. The fourth mentioned weakness in the cloud was the costs that have increased on the cloud side based on P4.

As a result of not having local implementation, there is a threat to lose potential customers in the long run based on P12 and P13.

"The customer see the data as a new oil...they do not want to share that data to others and especially to the competitors..." (P12)

## 6.2.2 Single edge strengths, weaknesses, opportunities and threats

When it comes to the new edge solution, there was already seen many **strengths** and opportunities. The strengths, weaknesses, opportunities, and threats of the single edge solutions based on the interviews are presented in Table 6. Six participants saw that the edge solution could make it easier for new customers to purchase digital solutions with lower risks since no data is transferred out of the site. Also, as P13 explained earlier the cloud makes it not transparent what data is transferred to the cloud, edge solution is different and more transparent for the customer. Based on P5, P9, and P10 the delay-sensitive IoT applications can benefit from the edge solution since there is less or non-existing latency. From an internal perspective, P1 saw it beneficial if the edge solution enables accessing second-level data instead of taking remote access to the DCS system in troubleshooting situations.

Table 6. *Single edge strengths, weaknesses, opportunities, and threats*

<b>Strengths</b>	<b>Weaknesses</b>	<b>Opportunities</b>	<b>Threats</b>
Lower threshold for the customer to purchase digital solutions	Complex to have access to the data	Easier to import offline data	Not everyone is familiar with remote access
Transparent what data is collected	Remote access has limitations: person count, time, personality	Decrease operational costs	If it is too difficult to access, people might not utilize the solution
Low latency	Low capacity		Maintenance
Suitable for accessing second-level data	Does not enable training more advanced analytical models		It is hard to get permission for remote access
			Can risk fleet monitoring and reference management capabilities

As an **opportunity**, P11 saw that it would be extremely beneficial if an edge solution could help in combining online and offline data since it has been difficult to import offline data to the cloud since it has been done manually. However, P6, P7, and P13 saw that the edge solution would make it even more difficult to combine online and offline data since analyzing online and offline data currently needs tools outside the platform of the case company.

"I think a valuable portion is to figure out how can we simplify the process when edge cloud isn't enough. ... I understand that customers don't want to give us data all the time but once they have a specific reason for giving us data, it will be super nice if the tool could enable us to sort of expand the use case in in different ways." (P13)

P1, P4, and P5 saw that edge solution could help in decreasing operational costs.

Similar to the cloud solution, also the edge solution had still some weaknesses and threats. The main **weakness** in the edge solution is the complexity of taking access to it based on eight participants. Another worry raised by P6, P7, and P11 is whether the remote access to edge has some limitations like the amount of people who can access the solution at the same time or time limitations. P6 underlined that these kinds of limitations will be fatal to remote services. P11 gave an example that sometimes there is a need to follow the test runs with remote access and it can last a whole working day and if there is a time limitation, this will not be possible. Another worry from P11 was that since the edge solution needs to be accessed with remote access, it will be difficult to replace someone during sick leave or vacation since the remote access rights are personal. Hence the remote service can suffer from this. When thinking about the capabilities, P17 stressed that when it comes to needs of high capacity, the edge equipment needs to be fierce to fulfill these requirements. Then the costs will also increase which is why the cloud is more suitable for these kinds of higher capacity requirements. P13 and P10 agreed with P17 and saw that the edge server does not enable the training of more advanced analytical models in the future.

"But I think it's very often that we need to do more with the data than just view the data with basic analytic tools inside the platform. So, I think it serves us some purposes, it serves us probably well for following up progress of agreements, doing basic type of analytics but I don't think it serves us enough to do advanced analytics." (P13)

One crucial **threat** with the edge solution is that some people in the case company have never used remote access and hence the edge solution would require a lot of training for this (P5, P15 & P16). P7, P11, and P13 said that if it is too difficult to take the remote

access, people might stop using it. P11 said that this has happened earlier as well with other new tools that people keep using the old ones if they are not fully pleased with the new tools. Another crucial threat to solve is how to maintain the edge since the maintenance is fully on the service provider and if the updates can not be done with remote access, all the updates need to be done one by one at site (P3; P10; P18).

"We have this SaaS and of course when we see bugs, we want to fix them in next release but it can not go so that the customer has to deal with the old version since we can not make the updates. It does not make sense." (P10)

P1 and P12 raised that even though customers do not want to transfer data to the cloud, it will get also more difficult to get remote access to the customer sites which then affects the availability of the digital solutions at the edge.

"So what concerns me is that the cases where we haven't had the cloud data, I think almost all of them have been that the customer has said like absolute no to any kind of remote access. So it doesn't help if it's an edge solution or cloud solution because there's not any kind of remote access allowed to decide. So we cannot still do remote monitoring." (P1)

P4 wondered whether the edge solutions ends up cutting the data transfer to the cloud completely which then cuts the possibility of fleet monitoring and proactive monitoring. P1 and P10 agreed that the edge solution can not risk the proactive monitoring aspect. Lastly, since customers are becoming even more careful, the edge needs to have the same security certificate as the DCS (P17).

### 6.2.3 DCS strengths, weaknesses, opportunities and threats

Even though the DCS system was not part of the scope in the interviews, it was raised in the discussions and hence the comments were included in this research to support the understanding of the overall picture. However, since DCS was not part of the interview scope, there are no comments from all participants. The strengths, weaknesses, opportunities, and strengths are presented in Table 7. Based on P8, P12, P14, and P15, the **strength** of the DCS is the possibility to get an overall picture of the process. It is possible to see all the measurements, how different controls are affecting each other, and what should be optimized. For P16 the DCS provides second-level data in troubleshooting cases that are not available in the cloud. P18 sees that it is nicer to use the customer tag names found from DCS when doing reporting since the tag names are then familiar to the customer. As a future **opportunity**, P17 saw that it would be beneficial to

identify which IoT applications should run in DCS instead of the cloud for example due to latency reasons.

Table 7. *DCS strengths, weaknesses, opportunities, and threats*

<b>Strengths</b>	<b>Weaknesses</b>	<b>Opportunities</b>	<b>Threats</b>
Get an overall understanding of the process	Not possible to combine online and offline data	Integrating IoT applications into DCS	It is becoming more difficult to get remote access
Access second-level data	Holds data only from a short period		
Tag naming is familiar for the customers			

As a **weakness**, the DCS system does not support combining online and offline data (P13) and it holds data only from the last half a year. In reference management, the data is required from a longer period (P12). As a threat, it will be more difficult to get remote access also to the DCS since the customers are becoming more careful with giving remote access (P12).

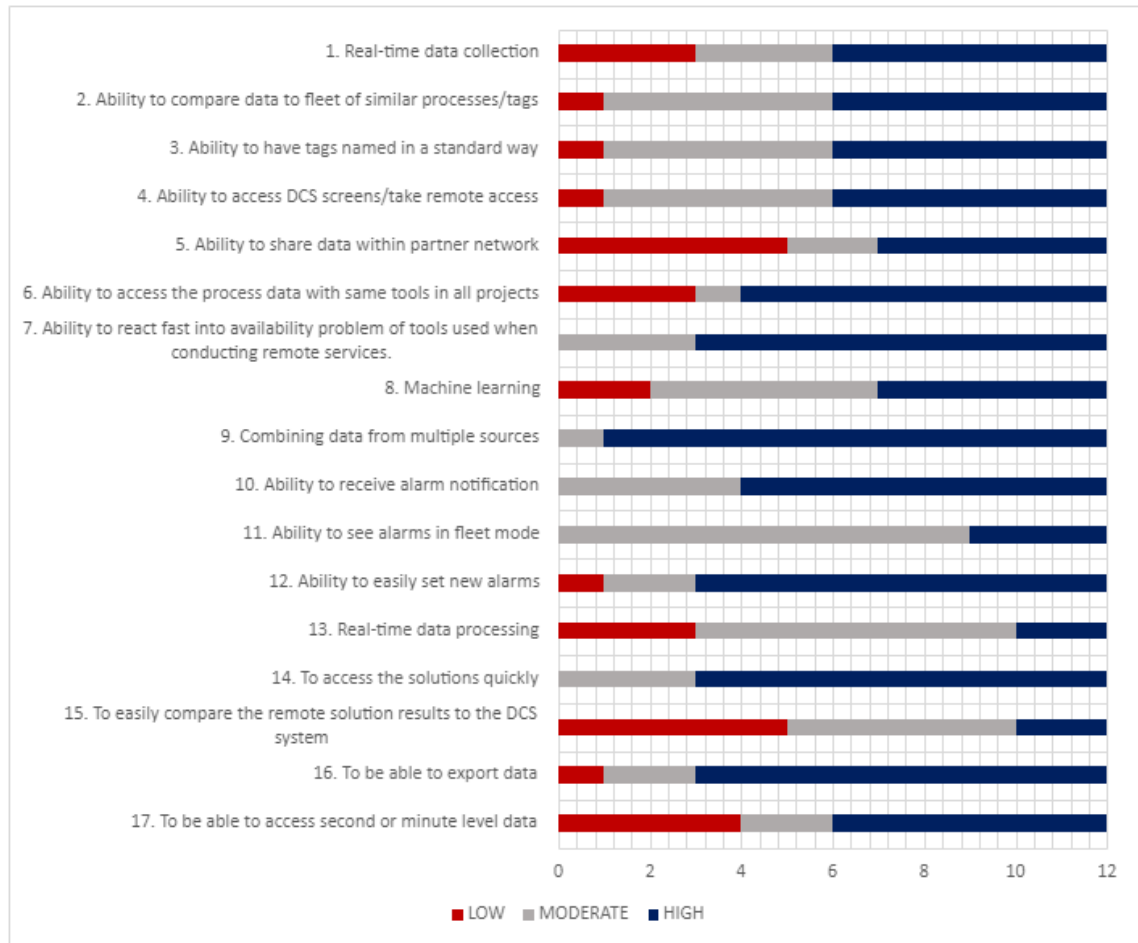
There were two reasons why it was seen that the cloud solution would be needed even though there would be a local solution. One was fleet monitoring and reference management (P1; P2; P5; P13; P14; P15). Based on P13 the cloud solution makes it possible to automate the work of an expert conducting remote services which would be lost in pure local implementation. The hybrid solution got a positive response from 13 out of 18 participants and the rest did not comment on the topic. The hybrid model is not in use yet, so it is not as clear for research participants. Hence the hybrid model was discussed more on a higher level and the aim was to understand how participants saw this possibility. The role of the edge in the hybrid model was seen in the pre-processing of data (P6; P10) which would then result in cost savings (P1; P3; P17) and provide transparency for the customer on what data is transferred to the cloud (P13).

#### **6.2.4 Critical capabilities for remote services**

Based on the questionnaire filled out during the interviews, there were seven common critical capabilities (2, 7, 9, 10, 12, 14, and 16 in Figure 15), eight capabilities with varying

needs (1, 3, 4, 5, 11, 13, 15 and 17 in Figure 15) and two nice to have capabilities (6 and 8 in Figure 15). 16 out of 18 participants saw that **fleet monitoring** has a lot of future potential, and that the interest towards it is increasing, one participant was unsure but thought that there might be potential in the future and one participant did not comment on the topic but saw it moderate in criticality. The major concern in fleet monitoring based on six participants is that the processes are difficult to compare since they are not standardized. Based on P10 and P13 some data, like energy consumption, are easier to compare than other more sensitive data. P8 and P11 highlighted that fleet monitoring requires a lot of data and hence P5 and P11 thought that it might not gain ground in a short time period. However, P10 and P14 said that there have already been some pilots on fleet monitoring but based on P10 there are not yet standardized processes. Based on P10 and P13 there needs to be a clear business model first. P9 and P14 see potential to utilize fleet data in troubleshooting, P14 sees that fleet monitoring should be part of inspections, and P5 and P12 see potential in reference management. P8, P11, P13, P14, and P17 also said that the customer hopes to monitor their fleet which needs to be enabled at some point in addition to the internal need. Five participants raised their worry about the possibilities of fleet monitoring when it comes to single edge solution.

**The ability to reach fast into availability problems** was seen as high or moderate in criticality by all participants. P5 and P8 see it as really disappointing to have to say to the customer that it was not possible to deliver the service due to problems with the tools. Based on P7 and P16 if the data is missing for over 50% of the reporting period, the service becomes useless. However, P15 and P16 see that it is okay if the data is not accessible in one or two days. Availability problems affect the time it takes to deliver the report to the customer and here even a few weeks of delay in delivering the report is too much (P11). This is why there should be better quality control and a clear process to define how the problem is identified, how the customer is contacted, and by who (P14). Cloud/remote connections are after all enabler for conducting remote services (P6).



**Figure 15.** Capabilities required for remote services.

Requirement of **combining data from different sources** was the highest rated requirement from all and P3 even said that it is probably the most important one from all since it helps to get the full understanding of the process. As an example, five participants mentioned the capability of combining process data with laboratory data. Other mentioned data were chemical measurement (P2) and maintenance data (P6). It was also said that it should be possible to combine data from 3rd party systems since the case company is also working with 3rd party DCS and then it might be more difficult to have access there (P3). One participant thought that the need to combine data from multiple sources might have been a reason to start providing cloud connection in the first place since it has not been possible locally (P10). Based on P16 data is already combined from multiple sources in the cloud which has been seen as useful. P4 raised that the capability of combining data from multiple sources is also important from the artificial intelligence perspective when determining what kind of data sources can be used to train the models. Earlier it was mentioned that P1, P4, and P10 are worried that the edge solution can risk the **proactive remote services** that are promoted and developed currently in the case company. P9 explained that the core of remote services is to take raw data and turn it

into an actionable item. Related to the proactive services, it was seen critical to have alarms and especially notifications from those so that the experts can react fast and proactively to monitored events. P15 explained that one expert is following multiple sites simultaneously and hence does not have time to monitor everything which is why there should be well-defined alarms that guide the work and where to focus. However, P1, P2, P5, and P13 stressed that the alarm notifications should not be automatic but instead it should be possible to set the notification only from the specified alarms and also the frequency of the notification since there are different types of alarms and not all of them need to send a notification, but they can be monitored other ways. P5 raised the importance of being able to modify existing alarms and their settings afterward since there can always be changes in processes that affect the alarm configuration. P10 and P16 saw that even though the customer has alarms already in their DCS system, these cloud-based alarms that are for long-term monitoring could also be sent to the customer if they are first validated by the case company experts so that no false alarms would be sent.

Another important capability related to these cloud-based alarms is that an expert can set new alarms themselves. This was rated as high in importance by 15 out of 18 participants. P13 rated this as low since there are already too many alarms for customers in DCS and it is difficult to sort the critical ones out of them all. However, P16 sees that the capability to set new alarms should not even be promoted for the customers but be an internal tool for proactive monitoring instead. P6 thought differently and saw that this capability would be important also for the customers. P3 and P6 saw this capability as pure user-friendliness since it actually takes time to make alarms with programming work based on P8.

It was seen as critical to **access the solutions quickly** by 15 out of 18 participants and it was said by P2 and P9 that if the solution is not accessed quickly, it will not be used either by the service provider or customer. P8 however raised delay as a more critical aspect than fast accessibility.

**Possibility to export data** was seen as important by 14 out of 18 participants. It was reasoned that data need to be exported if the tools are not good enough and the analysis need to be done with other tools (P2; P9). Another reason is that data is not in same place which is why it need to be exported and combined (P11). One participant said that instead of exporting raw data, it needs to be possible to export reports based on raw data (P8).

### 6.2.5 Different criticality of requirements inside the case company

The case company has different services that have different needs. Performance evaluation (reporting) does not require **real-time data** since it is done based on historical data and in intervals (P4; P7) and it is not about solving real-time issues (P2; P5; P11). However, the data that is used should be collected frequently and it should not have too many gaps (P8; P15). Also, the data delay should not be more than from hours to a day (P1; P11; P15; P16). Then when it comes to proactive monitoring, the data needs to be real-time data so that the events can be noticed and reacted to on time (P4; P6; P7; P14). Troubleshooting and some delay-sensitive applications also requires real-time data (P7; P9). In troubleshooting, if an issue is identified, it needs to be possible to verify it in real-time (P9). When it comes to delay-sensitive applications, if the data is not real-time, the information that the application provides is already useless (P10). P6 raised that not only case company experts should be interested in real-time data but also the service and even the customer.

**Standard tag naming** was seen as important but also difficult (P1; P3). P4 and P5 see that it would be straightforward and increase efficiency of work. P8 said that they already have standard tag naming but P5 said that they have been asking this for 15 to 20 years. P9 said that the tag naming is not as big of an issue if the descriptions are clear.

There are a lot of differences in the **utilization of DCS screens**. P1, P14, and P16 said that they have not needed this capability when doing remote services, P14 said that they don't even know how to use it but then P5, P10, and P18 said that they use remote connection all the time, and that it remains important in the future. P11 said that remote access is the main tool when customers have questions or problems. There were many reasons why remote access was seen as important: in troubleshooting to access second-level data (P3; P16; P18), cloud does not have all needed data (P4), DCS enables more analytical capabilities that are not in cloud (P3; P14; P18), DCS enables seeing how processes and machines are operated and how they are connected which is slower in cloud and tag descriptions in cloud are not clear enough (P5). Product teams were mentioned as one active user group for remote access (P1). P10 said that the cloud gives a first glance of the situation, but further analysis requires DCS. On the other hand, why remote access is not applicable every time is if the customer does not provide access (P9), sometimes DCS and cloud data are not comparable (P2) and remote access has restrictions like time that the user can spend there as mentioned earlier (P11). P15 said that even though they have not used remote access earlier, the utilization might increase in the future.

There is ongoing pilots about OEM partnership **ecosystems** based on P2, P4, and P14. Other than that, P9 and P16 mentioned that occasionally there is a need to do troubleshooting and collaboration with OEMs. P6 and P7 have not had a need yet for this kind of partnership. P8 said that the data is only collected from their own equipment which is why it will never be shared. P5 said that they see opportunities but there have not yet been business models that would benefit the case company, partner, and the customer. P4 said that it has been proved in one of the pilots that the case company can not do remote services for original equipment manufacturer (OEM) equipment as efficiently as OEM suppliers can and that the customer does not want to have multiple data connections but instead either their own or one from a supplier.

There are already existing use cases for **fleet-wide alarm monitoring** but it was seen as less important compared to the notification of an alarm by P4. P16 saw potential here for long-term monitoring purposes. P16 also sees that this could be used to contact the customer proactively about issues since sometimes problems are not communicated outside the operator room and the issues come as a surprise to the customer.

The answers related to **real-time data processing** were really different, but it can be explained by different needs in different tasks and cases. In long term monitoring and analysis based on historical data the delay was not seen as critical and data updates even once per day were seen okay (P3; P4; P9). But in emergency support and delay-sensitive applications like maintenance-related ones, the delay was seen as critical since too big of a delay makes the data invaluable (P3; P5; P10). P11 mentioned that data processing makes sometimes applications also slow which results in situations where the applications are not used.

**The possibility to compare data to the DCS system** was seen as low in criticality by 6 out of 18 participants and high by P10 and P18. P2 who voted for low said that data can not always be compared apple to apple. P11 who voted for low as well said that it should be possible to trust that the data is correct that is shown. However, P18 who voted for high said that it is common way of working in some cases that data is not showing correctly and it needs to be checked from the DCS.

When it comes to a need to **access second or minute-level data**, there was a lot of differences in answers. Once again it can be explained by different services with different needs. The less frequent data collection is enough when it comes to long term monitoring since there the historical data is analyzed and the focus is not in troubleshooting or finding issues but to focus on the long-term performance (P3). But when troubleshooting is

done where ad-hoc information, fast-developing events, or failures are analyzed, minute and second-level data are needed (P1; P2; P3; P5; P16).

**The ability to access all the process data always with the same tools** was seen as important from the efficiency of work perspective (P2; P7; P14) but after all that does not matter as long as the case company can deliver value to their customers (P3; P5). Same for **Machine Learning**, it is seen as an enabler for improved offering (P3; P4) but it is not value itself:

"I wouldn't say it's a value in self. ... if we need somewhere machine learning to identify for example the deviations of course but I don't say it's a goal to use machine learning. It's a goal to identify deviations." (P1)

## 7. DISCUSSION

This section goes through the research questions and answers those based on the literature review, questionnaire, and interview results. There were seven research questions that this thesis aimed to answer.

### 7.1 Remote services in process industry

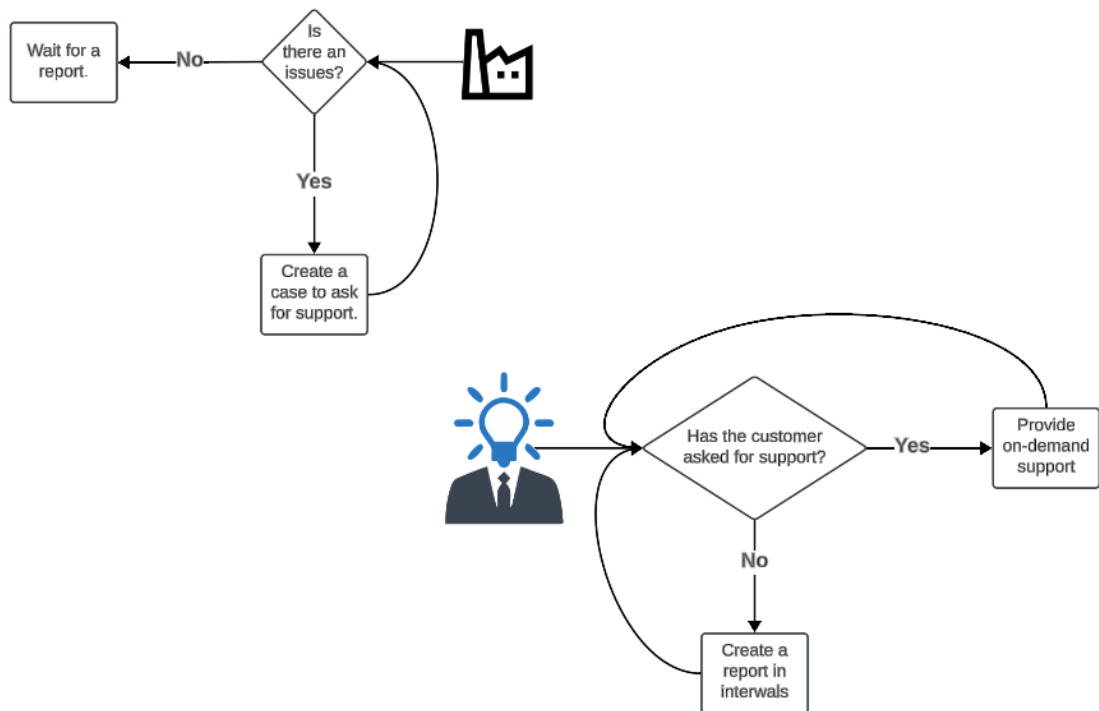
This section answers to RQ1. Since the process industry processes require continuous monitoring and optimization to prevent costly unplanned shutdowns, there is immense potential for service providers. Additionally, since IoT with its sensors and actuators has created data collection from the equipment, it enables a more versatile service offering where the service is provided based on data-based decision-making, data-driven services. The data collected from sensors is meaningless itself, but it becomes meaningful when it is processed, modeled, analyzed, visualized, evaluated, utilized in decision-making, and finally acted on as shown in Figure 6. This is an interesting perspective since many industrial customers tend to keep their data as a new oil. In fact, it becomes the oil only when the data is further processed and utilized somehow. All this data handling can be fully automated, especially in the process industry, where the automation system always exists, but most of the time to reach the wisdom level, human effort is needed. The human effort can be performed by the service provider, the customers themselves as a self-service, or in collaboration between the two as presented in Figure 7. In the case company, the services are provided mostly in collaboration with the customer, but the fully automated and customer self-service models might become more common in the future. This finding highlights the need for more advanced analytics.

A common way to provide data-driven services in the process industry is to transfer the collected data to the cloud and do remote services like remote monitoring. Cloud computing has created new opportunities like more advanced analytics and massive computing and storage capacity. This changes the traditional automation hierarchy model so that the upper-level, lower latency, functions can be migrated to the cloud as presented in Figure 10. Based on the case company, common services that are provided in the process industry are IoT applications, reporting services, and on-demand support services including 24/7 support. This finding was in line with the pre-understanding that existed.

As presented in Figure 4, there are four different types of analytics: descriptive, diagnostic, predictive, and prescriptive analytics. As defined earlier, the case company is in a chasm which means that the company has standardized ways to perform reactive analytics and services, but they do not yet have standardized processes for the proactive approach which requires predictive and prescriptive analytics. The analytical level focuses on descriptive and diagnostic analytics even though some more advanced and future-focused analytics are piloted. The overall service aims for a proactive approach but there is not yet standardized ways for this. The future direction is however to develop the analytical capabilities towards predictive and prescriptive capabilities and to create standardized processes for proactive service approach. Due to these findings, the maturity level of the analytics in the case company can be divided into the maturity of the services and the maturity of the technical capabilities to perform the services. Based on the findings it was surprising that the technical capabilities and the overall service were in different stages of maturity. The technical capabilities are in the chasm as the overall maturity of the analytics but the overall services are already in the mature and advanced/visionary stages aiming for a proactive approach.

The traditional way of working in remote services has been to report in intervals to the customer to provide support in identifying issues and optimization needs. This kind of monitoring and optimization is required in the process industry in order to perform well. Additionally, customers might have issues where they require support from the service provider outside the reporting intervals. In these situations, on-demand support is provided by the service provider. This traditional way of working is presented in Figure 16. However, the customer has started to ask for a more initiative-taking approach, proactive monitoring. Instead of reporting in intervals, the service provider would inform the customer whenever they see something alerting from the data:

” ...we have a problem with these our applications that we are continuously lacking. We look what has happened and it is extremely awkward that if we have a customer meeting twice a week and we tell something that has happened during the previous week that has no value... the customer knows about that themselves already.” (P10)



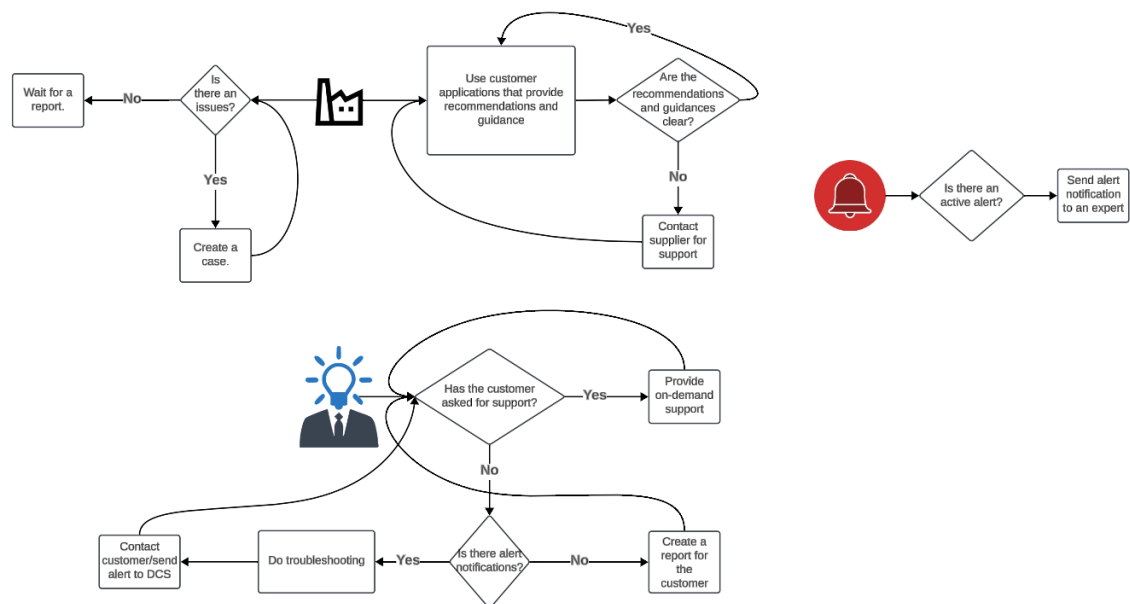
**Figure 16.** *Traditional way-of-working*

Alarms are commonly used to enable proactiveness so that a service provider or the customer gets information that can be acted on. Alarms are useful for the experts, in the company that provides the services, to guide their work. The experts are servicing multiple customers simultaneously and hence, they can not analyze everything, but they need an indication of where to focus. This is where these alarms can help so that the events are noticed as early as possible to gain business value:

"... It is like a accumulation function. It is like an expected value that we find something with diagnostics and the value increases until we do an action. And if the action is not done and the process is stopped, the value will drop to 0 ... then the information does not have value anymore... Not everything is this critical and usually the failures happen slower and we are able to act on it as long as we notice it." (P10)

This proactiveness and related alarms enable a new way of working which is presented in Figure 17. In addition to the customer contact or the reporting period, the service provider can proactively monitor the customer processes and contact the customer whenever something is found. The reporting can then provide information on the long-term

failures and optimization opportunities. The reporting can also include benchmarking services where the customer can get valuable insights that are based on anonymized customer data from the service provider installed base.



**Figure 17.** *Proactive way-of-working*

Another direction where the process industry is moving is to create more advanced analytics for the IoT applications. Then the applications can already give customers recommendations on optimal actions enabling customers to utilize the applications more in a self-service manner. This is presented in Figure 17 on top. Currently, IoT applications require human effort for evaluation:

”...they probably are deemed a bit too complex for customers to use just by themselves... So that customers would be able to create value by themselves of the application with just few support initiatives from us when needed, but today it's quite often so that they still rely on us to evaluate what the application is saying. ... And I think long term that's not a sustainable way to think that we can work with 200 and we will have to work hours with every customer” (P13)

In the process industry, the plant has multiple OEMs who know their own equipment the best. Since current OEM partnership pilot projects in the case company have shown that all OEMs have the best know-how from their equipment, ecosystem partnership models have been started. This can be seen as a future trend in the process industry even though the case company is one of the first runners in this kind of partnership model. In the ecosystem partnerships, the aim is to combine the OEM know-how to find conclusions and actions from data related to equipment to serve the common customer better. The customers have raised that they do not want to have multiple data connections for all OEMs separately. The ecosystem makes it possible to integrate all services into one

data connection and digital platform. Traditionally, the partnership between OEMs in the case company has been occasional. Sometimes solving a customer problem requires working together with OEMs. In these situations, collaboration has been done.

## 7.2 Tools and capabilities for performing remote services

This section answers to the RQ2. In the process industry, the customer has always an automation system (DCS). When the service provider is conducting the remote services, they must have remote access to the DCS so that they can get an overall understanding of the process, how controls are connected, and how the process is operated. In addition to remote access, there must be sensors that are connected to critical equipment for data collection. It is common nowadays that there is then a connection to the cloud for data transfer. The cloud enables conducting remote services without remote connection in an efficient way since the data is easier to access for analysis.

In the case company, the most utilized technical tools are cloud-based digital platform, customer portal, DCS screens, Python, commercial analytical tools, reporting tools like Power BI, and traditional Microsoft tools Excel and PowerPoint both for analysis and reporting purposes. The required tools and capabilities are summed in Table 8. It needs to be possible to visualize the data and create statistics from it, utilize the data in root cause analysis and regression analysis, and it needs to be possible to transfer the data into a reportable report. In the future, more advanced analytical models like machine learning models are required. The development is needed but there need to be also clear benefit from these more advanced models. In the case company, fleet monitoring capability was seen as potential and valuable in the future:

"It could be beneficial in a long run when we have a lot of data that we could compare similar equipment and processes in different plants... Currently when we talk about fleet, we mean that if the plant has multiple same type of asset, we look at them as a whole instead of focusing on them one by one to identify the worst performing ones." (P8)

Table 8. *Required tools and capabilities for performing remote services*

<b>Category</b>	<b>Item</b>	<b>Why is it needed?</b>
<i>Tool</i>	DCS screens	To get an overall picture of the process.

<i>Tool</i>	Digital platform	Where the data is visualized, post-processed, and analyzed.
<i>Tool</i>	Customer Portal	For delivering the IoT applications and the service for the customer. For communication between the service provider and the customer.
<i>Tool</i>	Reporting tools	To create a report that can be shared with the customer.
<i>Tool</i>	Additional analytical tools like commercial analytical tools and Python	For analysis purposes if the capabilities of the digital platform are not enough.
<i>Capability</i>	Data collection	Enabler for conducting remote services.
<i>Capability</i>	Access second-level data	Needed for troubleshooting.
<i>Capability</i>	Combine data from multiple sources	To get an overall understanding of the process and of maintenance actions and changes made in the process. Includes combining online and offline data.
<i>Capability</i>	Post-processing data	For analysis and to create reports.

<i>Capability</i>	More advanced analytical models	To increase the intelligence of the IoT applications and to enable proactive services based on future predictions.
<i>Capability</i>	Fleet monitoring	For efficiency of work (both service provider and customer), reference management, and for new kinds of services like benchmarking.
<i>Capability</i>	Alarm capabilities	To guide the work and for proactive services.
<i>Capability</i>	Easily accessed tools	For efficiency of work and to maintain users.
<i>Capability</i>	Self-service	To create content and make changes to existing content in an efficient and fast manner.
<i>Capability</i>	To export data	When additional analytical tools are needed where the data need to be transferred.

Fleet monitoring will be valuable for the service provider to guide where to focus, support inspection services, and in reference management. Fleet monitoring is usually seen as a tool for the service provider but it will also enable customers to monitor their fleet more efficiently. The customers are also interested in benchmarking services that could be integrated into traditional reporting practices. Not all key performance indicators (KPIs) are comparable since the data might be sensitive. There is still some KPIs that are comparable between customers that could be utilized in the benchmarking. It is important to identify all these different needs related to Fleet monitoring.

The tools need to be easy to access for efficiency of work and if it takes too long to access the tools or they work too slowly, no one will use them. This was a critical finding since the development focus of technical capabilities is usually on the features and not on efficiency of use. The user of the tools needs to be able to create their own content and make changes to existing content without programming work in a self-service manner. Otherwise, the development will take too long. The alarms themselves are important to move the services towards a proactive approach and to guide where the focus should be:

"When I start to have multiple customer plants where I do remote monitoring...it starts to require alarms that tell what should be looked at because it is impossible to monitor everything closely as it is required to be predictive." (P15)

Collected process data has different requirements. It does not have to be collected to the cloud in a second-level manner, but the second-level data needs to be accessible if needed in on-demand support. For reporting and on-demand support purposes, the delay is not critical since the focus is on historical data. The data that is collected can not contain too many gaps since then it is difficult to make the analysis based on it. In proactive monitoring, the delay is more critical since the focus is on real-time data. The tag naming in the data should be standard to make the work more efficient. There needs to be a possibility to combine online data with offline data. This is important in troubleshooting situations when all required data is not online. Offline data can provide critical information to understand the current situation. Additionally, there needs to be a possibility to combine data from multiple sources. The required data can reside in different sources but in an analysis situation, the data needs to be handled in one place.

### **7.3 Challenges of cloud**

This section answers to RQ3. As mentioned earlier, the IoT applications in the case company are too complex for the customers to analyze by themselves. This is why the reporting is needed to interpret the applications. The analysis requires expert work. Due to this, the service becomes expensive for the customer since they must pay for the application and the service on top of that. This is a crucial finding that creates a need for the future development of IoT applications. On the other hand, the applications can never replace the expertise that the case company has. Still, when the service becomes costly, it becomes more difficult to present value to the customer. Even though the applications are too complex, they still lack intelligence for example to do predictions. In the case company, there are not yet capabilities to do fleet monitoring although a lot of potential is seen there.

The biggest reason why the development of the edge solution has been started is the latency that the cloud has. The latency is caused by a large amount of data that is all the time increasing in the process industry. For delay-sensitive applications, this is fatal since it makes the results of the applications invaluable. Also, the latency makes the applications and tools slower which results in bad user experience. Hence, this finding is in line with the literature proposing that delay-sensitive applications should reside at the edge. Another driver for edge development is the increasing cost of data processing in the cloud. A third driver for edge is that customers are decreasingly willing to connect to the cloud and if there is no local solution available, it will result in losing potential customers. Similarly, customers are becoming more careful with providing remote access. In the future, these challenges can become more critical.

#### **7.4 Risks and opportunities of edge for remote services**

This chapter answers the research questions Q4 and Q5. Cloud computing is an efficient way of processing data but there is a need for edge computing. When the amount of data is growing, the cloud can not handle all tasks as efficiently, and timely as is required. This is a central finding in this research which highlights the need for the edge solution. Since the response time at the edge is from seconds to minutes, edge computing can help to reduce latency, network pressure, and bandwidth, which results in faster response times and better user experience due to better-performing tools. The fast response time is critical for delay-sensitive applications which is why they should be located at the edge. Edge can also support lower levels in automation hierarchy like controls and hence enable more automated operations. However, this kind of applications can also reside in DCS. The strength that the edge has over DCS is that edge can contain data from a longer period of time than DCS. On the other hand, DCS gives a better overall understanding of the process. Both can have second-level data but DCS can not have offline data available in addition to online data. Hence edge is more flexible than DCS in more advanced tasks.

Based on the literature review, one driver for the edge solution is the customer's need for better authority and desire to keep the data on-site. However, based on interviews the customers are not only unwilling to allow the cloud connection but also any kind of connection including remote connection which is required for conducting remote services with edge. Hence from a remote service perspective, the edge solution does not actually change this risk as much. Based on the interviews, increased data processing costs are also affecting customer willingness to transfer data to the cloud and they have surpris-

ingly even started to implement their own cloud solutions. It is hoped that the edge solution could decrease these data processing costs but there is not yet proof of that. Another reason why customers are unwilling to transfer their data to the cloud is security and privacy concerns. Edge can help in this concern since it is more transparent what data is residing in each location but on the other hand edge brings new security and privacy risks that need to be addressed and the edge solution must have the same security certificate as the DCS to appear appealing solution for the customers.

Another central finding in this research is that even though edge solution has many benefits and opportunities, it can not solve everything. Edge can not conduct tasks that require high capacity since otherwise, the edge equipment needs to include higher capacity which then increases the cost of investment. It is also harder to combine data from multiple sources or export data. These capabilities are required when performing remote services. It is also harder to maintain the edge solution if it needs to be done one by one. If the customer does not allow the updates through remote access the updates can not be done efficiently. Fleet monitoring is also not possible at the edge which is seen as an opportunity at the cloud side in the future. As remote access is critical to perform remote services with edge solution, it must be available for the service provider. However, it might be more difficult to get remote access in the future. The edge solution can not risk proactive monitoring. Hence, there needs to be alarm capabilities and alarm notification capabilities at the edge.

Edge computing is beneficial as an extension of the cloud, a hybrid model, instead of replacing it completely since the cloud still has many strengths and opportunities in the future. These strengths and opportunities are the possibility to combine data from multiple sources, fleet monitoring, and training more intelligent analytical models. Edge computing can be utilized to answer the low latency requirement and possibly to decrease data processing costs by conducting pre-processing. When the data is pre-processed in the edge, it can decrease the burden from the cloud and free more capacity for post-processing the data. Cloud can even be responsible for training more advanced analytical models which is not possible at the edge due to capacity issues. If the edge has high capacity, the cost of investment will also increase. However, the trained more advanced models can have a local version of them at the edge. The models can be trained in the cloud and transferred from there to the edge. This way edge and cloud can both benefit from each other for maximized performance and minimized costs as presented in Figure 11. However, in this kind of hybrid model, it is harder to define the business model. Another challenge is that to decrease data processing costs, not all data should be trans-

ferred to the cloud. On the other hand, if the cloud is used to train more intelligent analytical models, that requires a huge amount of high-quality data. The strengths, weaknesses, future opportunities, and threats of single cloud, single edge, hybrid, and DCS are summarized in the SWOT analysis in Figure 18.

	Cloud	Edge	Hybrid	DCS
Strengths	<ul style="list-style-type: none"> <li>- High computing and storage resources</li> <li>- More efficient work               <ul style="list-style-type: none"> <li>- Faster installed base monitoring</li> </ul> </li> <li>- Combining online and offline data</li> </ul>	<ul style="list-style-type: none"> <li>- Solution for customers that do not allow cloud connection</li> <li>- Better performance of applications</li> <li>- Low latency</li> </ul>	<ul style="list-style-type: none"> <li>- More transparency for the customer</li> <li>- Fleet monitoring capabilities</li> <li>- Better performance of applications</li> </ul>	<ul style="list-style-type: none"> <li>- Access second level data</li> <li>- See all measurements and how they are connected to each other               <ul style="list-style-type: none"> <li>- Get an overall picture of the process</li> </ul> </li> </ul>
Weaknesses	<ul style="list-style-type: none"> <li>- Contains latency</li> <li>- High data processing costs</li> </ul>	<ul style="list-style-type: none"> <li>- Slower installed base monitoring</li> <li>- Fleet monitoring not possible</li> <li>- More difficult to export data</li> <li>- More difficult to maintain</li> <li>- Investment becomes high with high computing and storage capacity</li> </ul>	<ul style="list-style-type: none"> <li>- Data in multiple environments</li> </ul>	<ul style="list-style-type: none"> <li>- Not offline data available</li> <li>- Includes data only for half a year period</li> </ul>
Opportunities	<ul style="list-style-type: none"> <li>- Combining data from multiple sources</li> <li>- Fleet monitoring and reference management</li> <li>- Training more intelligent models</li> </ul>	<ul style="list-style-type: none"> <li>- Broaden up service offering</li> <li>- Could decrease data processing costs</li> <li>- Could be connected to controls in DCS</li> <li>- Running local solutions from more intelligent models</li> </ul>	<ul style="list-style-type: none"> <li>- Possibility to decrease data processing costs</li> </ul>	<ul style="list-style-type: none"> <li>- Connecting delay-sensitive application to DCS</li> </ul>
Threats	<ul style="list-style-type: none"> <li>- Customers are increasingly unwilling to agree on cloud connection</li> </ul>	<ul style="list-style-type: none"> <li>- Remote services require remote connection</li> <li>- Harder to connect data from multiple sources</li> <li>- Decrease in proactive monitoring</li> <li>- No capacity for training more intelligent models</li> </ul>	<ul style="list-style-type: none"> <li>- What is the business model?</li> <li>- Not enough data for training more intelligent models</li> </ul>	<ul style="list-style-type: none"> <li>- Not enough data for training more intelligent models</li> </ul>

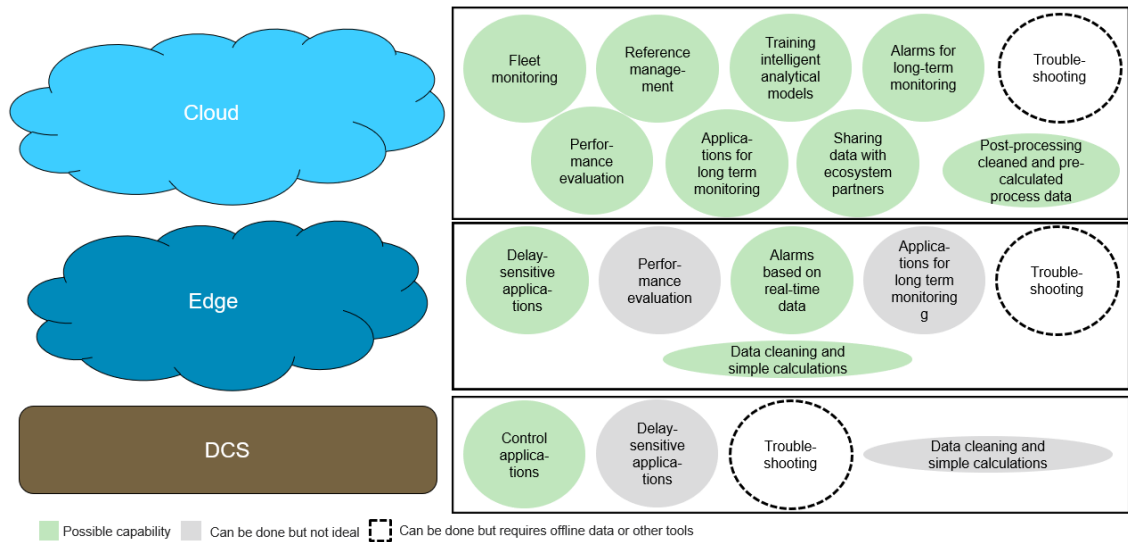
**Figure 18.** SWOT analysis of single cloud, single edge, hybrid solutions and DCS

Cloud, edge, and DCS have their roles and opportunities in the future. However, to make the most out of them, there needs to be a clear strategy on how to utilize them.

## 7.5 Role of edge in remote services

This section answers RQ6, and the resulting framework is the central finding from this research that summarizes the results. The edge solution enables that there can be local solutions from services and IoT applications that have traditionally resided in the cloud. However, the edge solution also provides an opportunity for a hybrid model where both local edge and decentralized cloud are used simultaneously. As presented in Figure 18, cloud, edge, and DCS have their own strengths. When having a hybrid model, there needs to be a decision on how all tasks are placed into the model. The strength of the cloud is the large resources and hence as presented in Figure 19, resource-oriented tasks like training intelligent analytical models and post-processing data should happen in the cloud. Another strength of the cloud is that it is easy to access the data when needed. Cloud enables also sharing data with ecosystem partners with customer permission. Fleet monitoring and reference management are also possible only in the cloud since they require combining data from multiple data sources and even different customers. Since the cloud has a delay, it can not answer the real-time requirement but it is still suitable for long-term monitoring. This is why IoT applications, performance evaluations

and alarms that do not require real-time data can reside in the cloud. Since they do not require real-time data, the data can be collected rarely, for example as daily aggregations to save costs. The collected data can also be pre-processed locally which reduces the data-processing costs at the cloud side. Troubleshooting can happen at the cloud but all data is not available there.



**Figure 19.** Task division between cloud, edge, and DCS

Since the edge is able to answer real-time requirements, delay-sensitive IoT applications and alarms that are based on real-time data should reside at the edge. Edge should also handle pre-processing data which can be data enlargement or data cleaning which can happen for example with an anomaly detection model. If there is only a local edge and no cloud, long-term monitoring, like performance evaluation, can also be done at the edge. Hence IoT applications and alarms that are utilized for long-term monitoring can also reside on edge. Troubleshooting is possible at the edge side. The edge provides a possibility to access second-level data for more precise analysis. However, when offline data is required, the edge does not support this. Finally, control applications should reside at DCS since they would be closer to the controls. DCS also has more data from the process than at the edge. Troubleshooting can be done at the DCS level but again, when offline data is needed, DCS does not support this.

The framework in Figure 19 provides generalized guidance on how edge, cloud, and DCS can be utilized together or separately, and what kind of capabilities each has. To perform remote services, there needs to be at least one of the three: DCS, edge, or cloud. However, if the cloud is not available, it is not possible to do fleet monitoring, reference management, or train more intelligent analytical models. If edge or DCS are available remote access is required so that it is possible to perform remote services. If

the edge or DCS is not available, it is not possible to have delay-sensitive IoT applications or alarms that utilize real-time data. DCS has alarms but these alarms need to be reacted immediately. When talking about the alarms in the edge or the ones in the cloud, they are more intelligent and not as critical as the ones in DCS.

In a hybrid model, the edge can be an extension of the cloud to answer real-time requirements and minimize the burden of the cloud and data processing costs. Edge can also provide local solutions from the cloud services and products if customers are unwilling to transfer their data to the cloud enabling a broader service and product portfolio. If the edge is utilized without the cloud, there needs to be an easy and efficient way for the service provider to access the data when needed to enable performing the service and proactive monitoring since it can not be done in the cloud. Also, then it is not possible or it is complicated to do fleet monitoring and reference management or to train more intelligent analytical models.

Even though cloud, edge, and DCS have their strengths, when placing services and digital products to each layer, there are still a few things that need to be considered. Costs, security and privacy, availability, scalability, accuracy, and user experience affect the placement decision as presented in Figure 12. As an example, if some application is resource-oriented, it would be best to place it into the cloud. However, if as a result the user experience is bad and the costs for data processing are high, it should be placed elsewhere or edge should be utilized at least in pre-processing of the data to decrease the costs and required bandwidth. Similarly, if an application that is delay-sensitive is placed to the edge but is not accurate due to data quality problems, it should be placed in DCS to ensure better data quality.

## **7.6 Ways to ensure performing remote services with edge solution**

This chapter answers the research question Q7. Things that need to be defined when developing an edge solution in the case company are presented in Table 9. From the perspective of pragmatism, these findings are the most important since these recommendations support concrete actions in the case company. Edge solution brings possibilities to minimize the burden of cloud and data processing costs. Edge can also broaden the service and product portfolio to serve customers who want a local solution but are still interested in the services and products on the cloud side. However, there are a few things to resolve before remote services can be performed with the edge solution efficiently. Remote access is critical to perform remote services and since the utilization of remote access is not standard in the case company, it should be trained. It should

always be included when the remote services are sold to the customer. Even though remote access is available, if it consists of restrictions like time or amount of people who can utilize it simultaneously, it makes it harder to perform remote services which is the current situation in the case company. Hence there should be an easy way to access the edge solution. A traditional way to take remote access to DCS is complicated, time-consuming, and contains restrictions. Since the edge does not contain all the data and information available, it could have its own, more simple, way to take remote access.

Table 9. *Things that need to be defined when developing an edge solution in the case company*

<b>Item</b>	<b>GAP in the case company</b>	<b>Required actions</b>
<b>Remote access</b>	<ul style="list-style-type: none"> <li>• Not all employees are familiar with taking remote access</li> <li>• Remote access is not always included when remote services are sold</li> <li>• Edge solution can be accessed with remote access but it is complicated and contains restrictions</li> </ul>	<ul style="list-style-type: none"> <li>• Training employees to utilize remote access</li> <li>• Remote access needs to be available for all employees who are conducting remote services and it needs to be included in the service contract</li> <li>• There needs to be an easier way to gain access to the edge solution</li> </ul>
<b>IoT applications at the edge</b>	Currently in the case company, there is no clear definition yet of which applications should be at the edge.	It needs to be defined which applications should be at the edge.
<b>Combining online and offline data</b>	Currently in the case company, there is no clear process for situations when online data at the edge	A process for combining online and offline data needs to be created.

	needs to be combined with offline data.	
<b>Customer portal and edge integration</b>	Customer portal in case company is an important communication channel between a service provider, customer, and partners. There is no local solution from the customer portal or integration with edge.	Integration between edge and customer portal needs to be created.
<b>Alarm notifications</b>	Currently, there is no way to send alarm notifications outside edge.	A way to send notifications outside edge from alarms needs to be created to not risk proactive monitoring.
<b>Relationship between cloud and edge</b>	Currently, there is no clear definition of the roles of edge and cloud if a hybrid model is utilized.	The roles of edge and cloud need to be defined.
<b>Maintenance of edge</b>	Currently, there is no clear plan on how to maintain the edge solution in the case company.	A well-defined maintenance plan needs to be created and needed resources need to be appointed.
<b>Availability of edge</b>	Currently, there are no alarms or other ways to follow the availability of the edge.	A way to monitor availability at the edge needs to be created.

Since edge solution is suitable for IoT applications that require real-time data, these kinds of applications need to be defined in the case company. There needs to be a clear understanding of what applications should be moved to edge and what kind of requirements

these bring to the capabilities of edge. Since edge does not have all the data that is required in troubleshooting, there needs to be a clear process for how a service provider is operating when the online data is not enough and offline data is needed. It also needs to be defined, what kind of tools can be utilized in this kind of situation. One option is to be able to export data from the edge to be able to combine the exported data with offline data with other tools. The possibility to export data is also important from the ecosystem point of view to share data with partners. A customer portal is an important tool from the communication point of view between the service provider, customer, and even ecosystem partners in the case company. One service in the case company is performed so that the analysis happens with analytical tools but the resulting recommendations are transferred to the customer through the customer portal. Hence there needs to be integration between the edge solution and the customer portal to ensure the knowledge transfer between different parties. To not risk proactive monitoring capabilities in the case company, the edge should enable sending alarm notifications for the service provider from monitored events. If a hybrid model is utilized, it is important to define the role of the edge in it.

There can be differences in whether a customer allows that edge solution can be updated and maintained through remote access or not. There needs to be a well-defined maintenance procedure for the edge solution and enough resources for this. The availability of the edge needs to be monitored and if issues are noticed it needs to be possible to react fast to those to maintain the solution as secure and up-to-date.

## 8. CONCLUSIONS

### 8.1 Answers to research questions

The **RQ1** studied what kind of remote services are provided in the process industry. Based on the findings from the case company, common services that are provided in the process industry are IoT applications, reporting services, and on-demand support services including 24/7 support. The services are usually performed as a collaboration between the service provider and the customer and the services are provided in intervals. The future direction is however to move to a more proactive approach where the service provider can react in real time to disturbances in processes. This requires predictive monitoring capabilities like alarms. Another future trend is to create more advanced analytics that can enable automated processes and customer self-service. Ecosystem partnerships will be important in the future to provide better services to customers.

The **RQ2** studied what technical capabilities and tools are required for conducting remote services. As presented in Table 8, the required tools based on the case company are a cloud-based digital platform, customer portal, DCS screens, Python, commercial analytical tools, reporting tools, and traditional Microsoft tools Excel and PowerPoint. To perform remote services there needs to be data collection, the possibility to access second-level data, the possibility to combine data from multiple sources, the possibility to pre-process data, advanced analytical models, fleet monitoring and alarming capability, the possibility to create own content as a self-service and to export data. Additionally, the tools need to be easily accessed to automate way of working. If the access is complicated or there are restrictions, the tools will not be utilized.

The **RQ3** studied what kind of shortages there are currently in capabilities in the cloud. Based on the findings, since the IoT applications are too complex, the customers can not use them as a self-service but then the service on top of the application makes it costly for the customer. Then it becomes harder to show the value of those. However, even though IoT applications are too complex for the customer, they still lack more advanced analytics. Also, advanced tools like fleet monitoring are currently missing in the case company. The edge solution is required for delay-sensitive IoT applications since the cloud has a delay. Another driver for the edge is the increased data processing costs. And thirdly, customers are becoming more careful and hence it will be harder to get customers to transfer their data to the cloud in the future.

The **RQ4** studied what kind of benefits edge solution could provide for remote services compared to a single cloud solution. As mentioned earlier in sub RQ3, edge computing can help to reduce latency and hence answer the needs that the delay-sensitive IoT applications have. Additionally, edge computing can decrease network pressure and bandwidth resulting in faster response times and better user experience. Edge computing can be utilized in lower levels of automation hierarchy, in controls, to automate operations. Edge computing provides a solution for those customers who want a local implementation and who do not want to transfer their data to the cloud due to low transparency, costs, or security and privacy concerns. In a hybrid model, the edge can act as an extension of the cloud by pre-processing data and holding delay-sensitive IoT applications. This way the cloud resources can be utilized more efficiently and for other tasks like training more advanced analytical models. It should be also studied further whether edge could hold local versions of more advanced analytical applications even though they would be trained in the cloud and whether edge could decrease costs in a hybrid model compared to the single cloud solution.

The **RQ5** studied what kind of risks and challenges there are in single edge solution. Edge can not have high capacity because then the up-front investment becomes high. Hence it can not train more advanced analytical models or hold IoT applications that require high capacity. Edge does not enable fleet monitoring. It is harder to combine data from multiple sources, export data, and maintain the solutions. Additionally, since remote access is critical for performing remote services with the edge, there is a risk that customers will not allow remote access as openly as today. This will then risk the proactive monitoring approach if a hybrid model is not in place and the proactive monitoring can not be performed in the cloud.

The **RQ6** studied how edge solution changes the way remote services are performed. Edge solution broaden ups the service and product portfolio to have local solutions. However, this requires that service provider need to train their employees to utilize remote access in their remote services. Cloud has its place when it comes to tasks that require high capacity, in fleet monitoring and reference management, combining data from multiple sources, and in ecosystem partnerships. Edge can answer real-time requirements and hold delay-sensitive IoT applications and alarms. Hence edge will be a tool for real-time monitoring and cloud will be a tool for long-term monitoring. The remote services can be performed with a single cloud, single edge, or a hybrid model. The decision is affected by costs, security and privacy, availability, scalability, accuracy, and user experience aspects.

The **RQ7** studied which things should be defined in the case company when developing edge solution further. Since remote access is not utilized by all in the case company it needs be trained. Since remote access is critical when performing remote services with the edge, it needs to be included in the contract every time the edge solution is sold. There needs to be easy access for the edge solution, the current way of taking access is too complicated. Things to develop are integration between the edge and the customer portal, alarm notification possibility outside edge, and a way to monitor the availability of the edge. Things to be defined are which IoT applications should reside at the edge, how to combine online and offline data, and what is the process for this, the roles of edge and cloud in a hybrid model, how the edge is maintained, and where to get resources for this.

## **8.2 Managerial implications**

The answer to the RQ7 gives implications for the case company on what things to focus on. One hypothesis in the case company was that the edge could replace the cloud and the cloud would no longer be needed. This hypothesis can be overruled based on findings that the cloud has many strengths and opportunities that the edge does not have. The long-term goals should be to have a clear business model of the utilization of edge and cloud which enables proactive monitoring and more advanced analytics. However, to reach this level, a lot of definitions need to be done on which IoT applications should reside at the edge, how to combine online and offline data and what is process for this, roles of edge and cloud in a hybrid model, how the edge is maintained and where to get resources for this. The base for these definitions should be the framework presented in Figure 19 which illustrates the strengths of cloud, edge, and DCS. Cloud should be utilized for long-term monitoring and for more advanced analytics while edge should be utilized for real-time analysis. There is not yet a trusted answer for whether the edge could help to decrease costs, which was another hypothesis in the case company. Based on this research it is possible, and hence it should be studied further. The important role of the DCS stays and in the future, it could be a place to automate processes with control applications.

In the meantime, it needs to be ensured that the edge solution can be utilized in remote services when a customer requests local solutions and that it does not risk performing remote services. The most critical thing is to enable easy access through remote access to the solution and training that increase the expertise of employees on the utilization of edge. The alarm notification is also an important feature to develop to continue the development of proactive monitoring. Similarly, integration to the customer portal from edge

is important to enable efficient knowledge transfer between service provider and customer.

### **8.2.1 Overall managerial recommendations**

Even though this research was done as a case study, the results can be generalized in the context of the process industry. Edge solution is a good addition to a portfolio of a service provider working in the process industry. Edge solution enables serving those customers who are more careful and do not want to transfer their data to the cloud but still are interested in the service and product portfolio on the cloud side. Additionally, edge helps to answer the real-time requirements even outside the process industry whenever latency on the cloud side is too much.

In a hybrid model, the edge can perform as an extension of the cloud or as an extension of DCS. Edge can help to reduce the burden of cloud computing and possibly decrease costs, and improve user experience even outside the process industry. When the edge is combined with DCS, it can help to bring more advanced analytics closer to DCS and help to automate control tasks in the process industry. All three, cloud, edge, and DCS have their strengths, weaknesses, opportunities, and threats but to reach the maximum benefit of remote services in the process industry, all of them are needed.

When developing an edge solution and remote services are connected to it, it is important to make sure that the capability to access the edge solution is not too complicated. Remote access should work efficiently without restrictions like time and amount of people who can access it simultaneously. Additionally, if a proactive monitoring approach is in use, the edge solution should support this instead of risking this. This can be done by enabling alarm notifications outside the edge or by utilizing a hybrid model where monitored data is transferred to the cloud after it is pre-processed at the edge. It is good to note that pre-processing is not only data cleaning but also data enlargement. If a customer portal is in use, the edge should be integrated into it to enable continuity of knowledge transfer between service provider and customer.

In the long run, there needs to be always a clear definition of the role of the edge, especially if a hybrid model is utilized: what solutions should reside at the edge, what are the processes if the capabilities at the edge are not enough, and how the solutions are maintained. This is valid for any kind of utilization of an edge solution.

This research provides a framework and guidance for companies in the process industry who want to develop an edge solution on its strengths, weaknesses, opportunities, and threats, how edge should be utilized in combination with cloud and DCS, and what things

need to be defined and enabled when developing an edge solution. This research provides guidance for all kind of companies that are interested in an edge solution of the strengths, weaknesses, opportunities and threats of the solution. Additionally, since there was a gap in the literature on how to take scalability, robustness, efficiency, manageability, and dynamicity into account when developing an edge solution, this research provides a better understanding of these. So that the solution can be scalable, robust, and manageable, there needs to be a clear process for maintaining it and a clear definition of what kind of applications and capabilities it holds. To be efficient, the edge solution should be utilized in line with its strengths instead of trying to accomplish everything at the edge. In the worst-case scenario, the up-front investment becomes high and the overall costs will increase due for example high energy consumption. To be dynamic, the edge solution should enable content creation as a self-service. From a dynamicity point of view, this research did not fully close the gap in the literature but it need to be further studied. Another gap in the literature was to consider how edge solution changes the way of performing remote services in the process industry. This is the main focus in this research and hence this gap is filled.

### 8.3 Evaluation of research

This research aimed for objectivity even though there were existing hypotheses and a background understanding of the topic and the case company. The results were a combination of theoretical insights and empirical results from a real-life case. The theory created a base that was then tested through a case study to add to the existing theory. The theory and empirical results were in line which is why the theory did not need adjustments but instead new frameworks were created on top of it.

It was decided to evaluate this research with Shenton's (2004) dimensions of comprehensive research which are credibility, transferability, reliability, and confirmability. **Credibility** refers to the extent to which the research results indicate the actual state of reality (Shenton 2004). Since edge computing is still a new technology, theory could not have given answers that would have indicated enough the state of reality. Hence, a case study was utilized as a support. The questionnaires and interviews were used in combination to get qualitative and quantitative data. The quantitative data made it possible to summarize things but the qualitative data created a deeper understanding of the reasons why people answered as they did. Also, the interviewer's background in the company and the pre-questionnaires created a good base understanding for the interviewer which made it possible to challenge the interviewees if the answers were not aligned with ex-

isting knowledge. However, even though an edge solution was presented to the interviewees, their existing knowledge and capability to understand the topic might have varied. Hence, the answers they provided were dependent on their own understanding of the topic. The participants were chosen with convenience sampling in a way that the sample would consist of people with different backgrounds and knowledge to get as wide sample as possible inside the scope of this research. However, these decisions affect the knowledge that was possible to be gathered.

**Transferability** means whether the results can be generalized and applied on a wider scale (Shenton 2004). Since this research was done as a case study, the results apply to similar companies working in the process industry and performing remote services. However, even though the overall research does not apply on a wider scale, parts of this research can be utilized everywhere where an edge solution is desired. This research can also work as a base for further research on the topic.

**Dependability** refers to the systematicity and transparency of the research and whether the research could be done again with the same results (Shenton 2004). In the literature review part, a clear plan was created and notes from all inquiries were made to keep track of searches and found materials. These notes were not presented in this document but with the presented search phrases, the most important materials will be found. In the empirical part, all answers were documented from the questionnaires and presented in this research. The interviews were transcribed and analyzed through those. The transcripts were not included as part of this document but the main findings are visible in the results and findings. Since the research process is presented well and in a transparent way in this document, the dependability is good and the research is repeatable.

Finally, **confirmability** refers to the objectivity of the research (Shenton 2004). As said earlier, this research aimed for objectivity. The pre-questionnaires and interviews were created with feedback from many people with different backgrounds. However, the hypothesis of the case company and existing knowledge might have still affected the content. The results were created individually but after that multiple feedback rounds were held to refine the content so that it is realistic and fits the real-life context. These feedback rounds gave new perspectives and reduced biases.

The combination of theoretical insights and empirical results from the case study provided results that are both applicable to real-life contexts and generalizable. The biases were minimized with multiple feedback loops and by transparent documentation of the research design and methods.

## 8.4 Future work

This research created a base understanding of the edge solution, its characteristics, and utilization opportunities. However, since the focus was on the process industry, the research could be scaled for another context as well.

One hypothesis that existed in the case company was whether edge computing could be utilized to minimize costs that have increased in the cloud. This research failed to address this hypothesis since the scope of this research would have gone too wide. Hence this is an important topic to study further. This will require studying a real-life case and comparing the costs over a longer period to get accurate results. Based on this research it is proved that there is potential for savings in costs with edge computing but it needs to be verified.

This research does not focus either on the technical implementation of the edge solution which should be studied in the future. In this research, it was identified that there can be many different ways to implement an edge solution which is why it needs to be further studied to identify the optimal way in the process industry context. Fog computing was briefly mentioned in this research but it was left out of the scope of the research since the technical implementation was not in the focus. However, as shown in this research, there are many different definitions for fog computing, hence the definition of fog computing needs to be standardized. Additionally, the utilization opportunities of fog computing should be identified.

There were gaps in the literature about edge computing which were addressed in this research. However, the dynamicity aspect remained narrow and hence it should be further studied. Dynamicity is an important aspect of the edge. Currently, the dynamicity at the cloud is not optimal and hence the same shortage should not exist at the edge.

## REFERENCES

- Ackoff, R. L. (1989). From Data to Wisdom. *Journal of Applied Systems Analysis*. Vol. 16(1). pp. 3-9.
- Angelopoulos, A., Michailidis, E.T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S. & Zahariadis, T. (2019). Tackling Faults in the Industry 4.0 Era—A Survey of Machine-Learning Solutions and Key Aspects. *Sensors* 2020. Vol. 20(109). <https://doi.org/10.3390/s20010109>.
- Ashouri, M., Davidsson, P. & Spalazzese, R. (2018). Cloud, Edge, or Both? Towards Decision Support for Designing IoT Applications. *Fifth International Conference on Internet of Things: Systems, Management and Security*. pp. 155–162. <https://doi.org/10.1109/IoTSMS.2018.8554827>.
- Azeved, R. (2022). Predict Process Events with Contextual Data Insights: Maintenance Teams See Predicted Machine Failures. *Plant Engineering*. Available at: <https://www.plantengineering.com/articles/predict-process-events-with-contextual-data-insights/>. (Read: 18.6.2024)
- Bellavista, P., Penna, R. D., Foschini, L. & Scotece, D. (2020). Machine Learning for Predictive Diagnostics at the Edge: An IIoT Practical Example. *IEEE International Conference on Communications (ICC)*. pp. 1–7. <https://doi.org/10.1109/ICC40277.2020.9148684>.
- Bollapragada, S., Gupta, A. & Lawsirirat, C. (2007). Managing a portfolio of long term service agreements. *European Journal of Operational Research*. Vol. 182. pp. 1399–1411. <https://doi.org/10.1016/j.ejor.2006.08.046>
- Booth, A. (2015). Unpacking your literature search toolbox: On search styles and tactics. *Health Information & Libraries Journal*. DOI: 10.1111/j.1471-1842.2008.00825.x
- Colombo, A. W., Karnouskos, S., Kaynak, O., Shi, Y. & Yin, S. (2017). Industrial Cyber-physical Systems: A Backbone of the Fourth Industrial Revolution. *IEEE Industrial Electronics Magazine*. Vol. 11(1). pp. 6–16. <https://doi.org/10.1109/MIE.2017.2648857>.
- Dai, W., Zhang, Y., Kong, L., Christensen, J. H. & Huang, D. (2023). Design of Industrial Edge Applications Based on IEC 61499 Microservices and Containers. *IEEE Transactions on Industrial Informatics*. Vol. 19(7). pp. 7925–35. <https://doi.org/10.1109/TII.2022.3214199>.
- Dastjerdi, A. V. & Buyya, R. (2016). Fog Computing: Helping the Internet of Things Realize Its Potential. *Computer*. Vol. 49(8). pp. 112–16. <https://doi.org/10.1109/MC.2016.245>.

- Eigner, A. & Stary, C. (2023). The Role of Internet-of-Things for Service Transformation. *SAGE Open*. Vol. 13(1). <https://doi.org/10.1177/21582440231159281>.
- El-Sayed, H., Sankar, S., Prasad, M., Puthal, D., Gupta, A., Mohanty, M. & Lin, C-T. (2018). Edge of Things: The Big Picture on the Integration of Edge, IoT and the Cloud in a Distributed Computing Environment. *IEEE Access*. Vol. 6. pp. 1706-1717. <https://doi.org/10.1109/ACCESS.2017.2780087>.
- Francis, T. (2018). A Comparison of Cloud Execution Mechanisms Fog, Edge, and Clone Cloud Computing. *International Journal of Electrical and Computer Engineering (IJECE)*. Vol. 8(6). <https://doi.org/10.11591/ijece.v8i6.pp4646-4653>.
- Gebauer, H., Paiola, M., Sacconi, N. & Rapaccini, M. (2021). Digital Servitization: Crossing the Perspectives of Digitization and Servitization. *Industrial Marketing Management*. Vol. 93. pp. 382–88. <https://doi.org/10.1016/j.indmarman.2020.05.011>.
- Grubic, T. (2018). Remote Monitoring Technology and Servitization: Exploring the Relationship. *Computers in Industry*. Vol. 100. pp. 148–158. <https://doi.org/10.1016/j.compind.2018.05.002>.
- Hafeez, T., Xu, L. & Mcardle, G. (2021). Edge Intelligence for Data Handling and Predictive Maintenance in IIOT. *IEEE Access*. Vol. 9. pp. 49355-49371. <https://doi.org/10.1109/ACCESS.2021.3069137>.
- Halper, F. (2020). TDWI Analytics Maturity Model: Assessment Guide. Transforming Data With Intelligence. Available at: <https://tdwi.org/pages/assessments/adv-all-tdwi-analytics-maturity-model-assessment.aspx>. (Read: 18.6.2024)
- Hassan, N., Gillani, S., Ahmed, E., Yaqoob, I. & Imran, M. (2018). The Role of Edge Computing in Internet of Things. *IEEE Communications Magazine*. Vol. 56(11). pp. 110–115. <https://doi.org/10.1109/MCOM.2018.1700906>.
- Hästbacka, D., Halme, J., Barna, L., Hoikka, H., Pettinen, H., Larranaga, M., Bjorkbom, M., Mesia, H., Jaatinen, A. & Elo, M. (2022). Dynamic Edge and Cloud Service Integration for Industrial IoT and Production Monitoring Applications of Industrial Cyber-Physical Systems. *IEEE Transactions on Industrial Informatics*. Vol. 18(1). pp. 498–508. <https://doi.org/10.1109/TII.2021.3071509>.
- Huang, H., Yang, L., Wang, Y., Xu, X. & Lu, Y. (2021). Digital Twin-Driven Online Anomaly Detection for an Automation System Based on Edge Intelligence. *Journal of Manufacturing Systems*. Vol. 59. pp. 138–150. <https://doi.org/10.1016/j.jmsy.2021.02.010>.
- ISA–95.00.01–CDV3. (2008). Enterprise-Control System Integration Part 1: Models and Terminology. North Carolina: Research Triangle Park. 2008. 165 s.

- Isaksson, A. J., Harjunkoski, I. & Sand, G. (2018). The Impact of Digitalization on the Future of Control and Operations. *Computers & Chemical Engineering*. Vol. 114. pp. 122–129. <https://doi.org/10.1016/j.compchemeng.2017.10.037>.
- ISO 13372:2012. (2012). Condition monitoring and diagnostics of machines — Vocabulary. Edition 2. Technical Committee ISO/TC. 25 s.
- Jin, J., Yu, K., Kua, J., Zhang, N., Pang, Z. & Han, Q-L. (2024). Cloud-Fog Automation: Vision, Enabling Technologies, and Future Research Directions. *IEEE Transactions on Industrial Informatics*. Vol. 20(2). pp. 1039–1054. <https://doi.org/10.1109/TII.2023.3272696>.
- Juergen, J. & Givehchi, O. (2013). Industrial Automation Services as Part of the Cloud: First Experiences. Available at: [https://www.researchgate.net/publication/257402460\\_Industrial\\_Automation\\_Services\\_as\\_part\\_of\\_the\\_Cloud\\_First\\_experiences](https://www.researchgate.net/publication/257402460_Industrial_Automation_Services_as_part_of_the_Cloud_First_experiences). (Read: 18.6.2024)
- Jussila, J., Suominen, A.H., Partanen, A., Honkanen, T. (2021). Text Analysis Methods for Misinformation–Related Research on Finnish Language Twitter. *Future Internet*. Vol. 13(157). <https://doi.org/10.3390/fi13060157>
- Järvelin, K. & Sormunen, E. (2010). Tiedon tallennus ja haku. Ote informaatiosta: johdatus informaatiotutkimukseen ja interaktiiviseen mediaan. pp. 155-210.
- Kanninen, T., Penttinen, E., Tinnilä, M. & Kaario, K. (2017). Exploring the Dynamic Capabilities Required for Servitization: The Case Process Industry. *Business Process Management Journal*. Vol. 23(2). pp. 226–247. <https://doi.org/10.1108/BPMJ-03-2015-0036>.
- 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*. Vol. 56(2). pp. 44–51. <https://doi.org/10.1109/MCOM.2018.1700622>.
- Kong, L., Tan, J., Huang, J., Chen, G., Wang, S., Jin, X., Zeng, P., Khan, M. & Das, S. K. (2023). Edge-Computing-Driven Internet of Things: A Survey. *ACM Computing Surveys*. Vol. 55(8). pp. 1–41. <https://doi.org/10.1145/3555308>.
- Król, K. & Zdonek, D. (2020). Analytics Maturity Models: An Overview. *Information*. Vol. 11(3):142. <https://doi.org/10.3390/info11030142>.
- Krumeich, J., Werth, D. & Loos, P. (2016). Prescriptive Control of Business Processes: New Potentials Through Predictive Analytics of Big Data in the Process Manufacturing Industry. *Business & Information Systems Engineering*. Vol. 58(4). pp. 261–280. <https://doi.org/10.1007/s12599-015-0412-2>.

- Langmann, R. & L. Meyer. (2014). Automation Services from the Cloud. 11th International Conference on Remote Engineering and Virtual Instrumentation (REV). pp. 256–261. <https://doi.org/10.1109/REV.2014.6784271>.
- Liu, B., Zhang, Y., Zhang, G. & Zheng, P. (2019). Edge-Cloud Orchestration Driven Industrial Smart Product-Service Systems Solution Design Based on CPS and IIoT. *Advanced Engineering Informatics*. Vol. 42. <https://doi.org/10.1016/j.aei.2019.100984>.
- Loghin, D., Ramapantulu, L. & Teo, Y. M. (2019). Towards Analyzing the Performance of Hybrid Edge-Cloud Processing. In 2019 IEEE International Conference on Edge Computing (EDGE). pp. 87–94. <https://doi.org/10.1109/EDGE.2019.00029>.
- Mubeen, S., Nikolaidis, P., Didic, A., Pei-Breivold, H., Sandstrom, K. & Behnam, M. (2017). Delay Mitigation in Offloaded Cloud Controllers in Industrial IoT. *IEEE Access*. Vol. 5. pp. 4418–4430. <https://doi.org/10.1109/ACCESS.2017.2682499>.
- Ning, H., Li, Y., Shi, F. & Yang, L. T. (2020). Heterogeneous Edge Computing Open Platforms and Tools for Internet of Things. *Future Generation Computer Systems*. Vol. 106. pp. 67–76. <https://doi.org/10.1016/j.future.2019.12.036>.
- Ometov, A., Molua, O., Komarov, M. & Nurmi, J. (2022). A Survey of Security in Cloud, Edge, and Fog Computing. *Sensors*. Vol. 22(3):927. <https://doi.org/10.3390/s22030927>.
- Pan, J. & McElhannon, J. (2018). Future Edge Cloud and Edge Computing for Internet of Things Applications. *IEEE Internet of Things Journal*. Vol. 5(1). pp. 439–49. <https://doi.org/10.1109/JIOT.2017.2767608>.
- Pearson, S. & Benameur, A. (2010). Privacy, Security and Trust Issues Arising from Cloud Computing. In 2010 IEEE Second International Conference on Cloud Computing Technology and Science. pp. 693–702. <https://doi.org/10.1109/CloudCom.2010.66>.
- Qian, G., Lu, S., Pan, D., Tang, H., Liu, Y. & Wang, Q. (2019). Edge Computing: A Promising Framework for Real-Time Fault Diagnosis and Dynamic Control of Rotating Machines Using Multi-Sensor Data. *IEEE Sensors Journal*. Vol. 19(11). pp. 4211–4220. <https://doi.org/10.1109/JSEN.2019.2899396>.
- Raddats, C., Naik, P. & Bigdeli, A. Z. (2022). Creating Value in Servitization through Digital Service Innovations. *Industrial Marketing Management*. Vol. 104. pp. 1–13. <https://doi.org/10.1016/j.indmarman.2022.04.002>.
- Raman, R. & Caeron, I. (2005). Abnormal Situation Management in Process Control: Do we have it under control? 7th World Congress of Chemical Engineering. pp. 10-14.

- Rapaccini, M. and Adrodegari, F. (2022). Conceptualizing Customer Value in Data-Driven Services and Smart PSS. *Computers in Industry*. Vol. 137:103607. <https://doi.org/10.1016/j.compind.2022.103607>.
- Ren, L., Jia, Z., Wang, T., Ma, Y. & Wang, L. (2022). LM-CNN: A Cloud-Edge Collaborative Method for Adaptive Fault Diagnosis With Label Sampling Space Enlarging. *IEEE Transactions on Industrial Informatics*. Vol. 18(12). pp. 9057–9067. <https://doi.org/10.1109/TII.2022.3180389>.
- Saunders, M. N. K., Lewis, P. & Thornhill, A. (2007). *Research Methods for Business Students*. 8th ed. Pearson Education. Harlow, England.
- Schmidt, B., Borrison, R., Cohen, A., Dix, M., Gärtler, M., Hollender, M., Klöpffer, B., Maczey, S. & Siddharthan, S. (2018). Industrial Virtual Assistants: Challenges and Opportunities. *ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. pp. 794–801. <https://doi.org/10.1145/3267305.3274131>.
- Shenton, A. K. (2004). Strategies for Ensuring Trustworthiness in Qualitative Research Projects. *Education for Information*. Vol. 22(2). pp. 63–75. <https://doi.org/10.3233/EFI2004-22201>
- Shi, W., Cao, J., Zhang, Q., Li, Y. & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*. Vol. 3(5). pp. 637–646. <https://doi.org/10.1109/JIOT.2016.2579198>.
- Suter, M., Eidenbenz, R., Pignolet, Y-A. & Singla, A. (2019). Fog Application Allocation for Automation Systems. *IEEE International Conference on Fog Computing (ICFC)*. pp. 97–106. <https://doi.org/10.1109/ICFC.2019.00021>.
- Taneja, M. & Davy, A. (2017). Resource Aware Placement of IoT Application Modules in Fog-Cloud Computing Paradigm. *IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*. pp. 1222–1228. <https://doi.org/10.23919/INM.2017.7987464>.
- Teoh, Y. K., Gill, S. S. & Parlikad, A. K. (2023). IoT and Fog-Computing-Based Predictive Maintenance Model for Effective Asset Management in Industry 4.0 Using Machine Learning. *IEEE Internet of Things Journal*. Vol. 10(3). pp. 2087–2094. <https://doi.org/10.1109/JIOT.2021.3050441>.
- Tusa, F., Clayman, S., Buzachis, A. & Fazio, M. (2024). Microservices and Serverless Functions—Lifecycle, Performance, and Resource Utilisation of Edge Based Real-Time IoT Analytics. *Future Generation Computer Systems*. Vol. 155. pp. 204–218. <https://doi.org/10.1016/j.future.2024.02.006>.
- Verdugo-Cedeño, M., Jaiswal, S., Ojanen, V., Hannola, L. & Mikkola, A. (2023). Simulation-Based Digital Twins Enabling Smart Services for Machine Operations: An

- Industry 5.0 Approach. *International Journal of Human–Computer Interaction*. pp. 1–17. <https://doi.org/10.1080/10447318.2023.2254607>.
- Wang, T., Ke, H., Zheng, X., Wang, K., Sangaiah, A. K. & Liu, A. (2020). Big Data Cleaning Based on Mobile Edge Computing in Industrial Sensor-Cloud. *IEEE Transactions on Industrial Informatics*. Vol. 16(2). pp. 1321–1329. <https://doi.org/10.1109/TII.2019.2938861>.
- Wohlin, C. (2014). Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. 18th International Conference on Evaluation and Assessment in Software Engineering. pp. 1–10. DOI: 10.1145/2601248.2601268.
- Wolniak, R. & Grebski, W. (2023). THE FIVE STAGES OF BUSINESS ANALYTICS. Silesian University of Technology, Organization and Management Department, Economics and Informatics Institute. <https://doi.org/10.29119/1641-3466.2023.178.41>
- Xu, H., Yu, W., Griffith, D. & Golmie, N. (2018). A Survey on Industrial Internet of Things: A Cyber-Physical Systems Perspective. *IEEE Access*. Vol. 6. pp. 78238–78259. <https://doi.org/10.1109/ACCESS.2018.2884906>.
- Zheng, P., Wang, Z., Chen, C-H. & Khoo, L. P. (2019). A Survey of Smart Product-Service Systems: Key Aspects, Challenges and Future Perspectives. *Advanced Engineering Informatics*. Vol. 42: 100973. <https://doi.org/10.1016/j.aei.2019.100973>.

# APPENDIX A: QUESTIONNAIRE FILLED IN THE INTERVIEWS

## 1. Interviewee \*

Enter your answer

## 2. Business Line \*

Enter your answer

## 3. How important do you see these capabilities from your remote service point of view? (Now-2years) \*

	Low	Moderate	High
Real-time data collection	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to compare data to fleet of similar processes/tags	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to have tags named in a standard way	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to access DCS screens/take remote access	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to share data within partner network	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to access the process data with same tools in all projects	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to react fast into availability problem of tools used when conducting remote services.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Machine learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Combining data from multiple sources	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to receive alert notification from cloud alerts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to see cloud alerts in fleet mode	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ability to easily set new cloud alerts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Real-time data processing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To access the solutions quickly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

To easily compare the remote solution results to the DCS system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To be able to export data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To be able to access second or minute level data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## APPENDIX B: PRE-QUESTIONNAIRE ANSWERS

There were additional background questions and questions for case company purposes that were not utilized in this research. The name of the digital platform in the case company is hidden from question five.

3. What kind of remote services are you offering? (Select one or more)

[More Details](#)

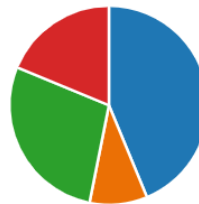
<span style="color: blue;">●</span> Industrial Internet Applications	12
<span style="color: orange;">●</span> Reporting	13
<span style="color: green;">●</span> Troubleshooting / on-demand s...	13
<span style="color: red;">●</span> 24/7 support	3
<span style="color: purple;">●</span> Other	2



4. Who is conducting the service? e.g. first technology, then internal expert (Select one or more)

[More Details](#)

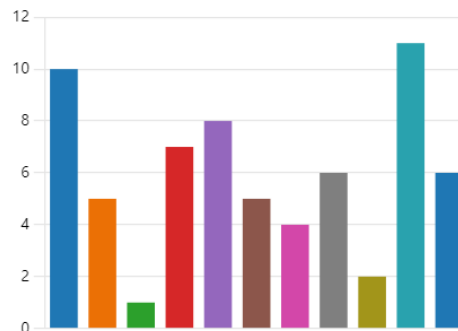
<span style="color: blue;">●</span> Internal expert	14
<span style="color: orange;">●</span> Customer as self-service	3
<span style="color: green;">●</span> Collaboration between internal ...	9
<span style="color: red;">●</span> Technology (e.g. application)	6



5. What kind of tools are currently critical for conducting these remote services? (Select one or more)

[More Details](#)

<span style="color: blue;">●</span> [Tool]	10
<span style="color: orange;">●</span> Tableau	5
<span style="color: green;">●</span> Power BI	1
<span style="color: red;">●</span> Customer Portal	7
<span style="color: purple;">●</span> Excel	8
<span style="color: brown;">●</span> Salesforce	5
<span style="color: pink;">●</span> Snowflake	4
<span style="color: grey;">●</span> Power Point	6
<span style="color: olive;">●</span> Word	2
<span style="color: teal;">●</span> DCS screens	11
<span style="color: blue;">●</span> Other	6



6. Which questions does your service answer? (Select one or more)

[More Details](#)

- What has happened? 11
- Why something happened? 14
- What will/might happen? 12
- What should be done to preven... 15



7. What is/are the focus in your service? (Select one or more)

[More Details](#)

- Detect future trends & possibilit... 5
- Perform root cause analysis 11
- Define past performance & tren... 12
- Define best course of action 8



8. What is the time perspective in your service?

[More Details](#)

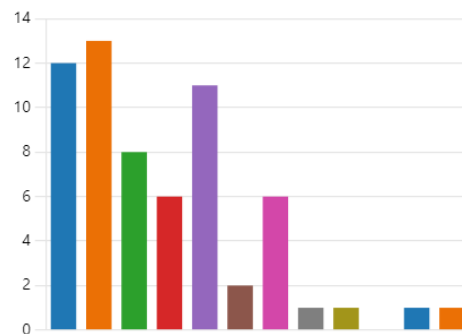
- Historical data (analysis is based... 15
- Future predictions (historical dat... 6



9. Which techniques/models are used in your service? (Select one or more)

[More Details](#)

- Data visualizations 12
- Reporting 13
- Statistics 8
- Drill-down analysis (shift from o... 6
- Root cause analysis 11
- Machine learning 2
- Regression analysis (estimate rel... 6
- Forecasting 1
- Decision trees (if x happens the... 1
- Simulation models 0
- Optimization models 1
- Other 1



10. How does your service support decision-making? (Select one or more)

[More Details](#)

- Give information 14
- Anticipate risks and/or opportu... 7
- Guide optimal decision to achie... 7
- Identify areas for improvement ... 12



14. In some cases the customer are unwilling to transfer data to cloud. Do you see that edge solution could boost our chances of winning potential sales cases?

[More Details](#)

- Yes 5
- Sometimes 4
- I'm not sure 5
- No 1

