Data driven decision making in Industrial Control Systems

Andreas Bäck

Degree Thesis for Master of Engineering

Degree Program – Technology Based Management

Vaasa 2019

EXAMENSARBETE

Författare:	Andreas Bäck
Utbildningsprogram och ort:	Teknologibaserat Ledarskap, Vasa
Handledare:	Stefan Granqvist, Yrkeshögskolan Novia
	Leif Strandberg, Wärtsilä

Titel: Data driven decision making in Industrial Control Systems

---

Datum 12.12.2019   Sidantal 58

Abstrakt

Denna avhandling handlar om att se nyttan och värdet med att ha en datadriven utveckling ifråga om automationssystem som finns i 4 - takts motorsystem i Wärtsilä Finland Oy.

Det främsta syftet med avhandlingen är att dels beskriva och förklara men även att *ge råd* om hur data kan användas för att kunna göra ännu bättre och lönsamma beslut för industriella kontrollsystem. Det andra syftet är att ta fram och presentera verktyg och modeller för att med data kunna nå målen.

Avhandlingen genomfördes genom inventering av litteratur och forskning om hur man skall gå till väga för att transformera till en mera data driven business. Studien ska ses som en aktionsanalytisk studie där eget engagemang, kurser och företagsintern dialog inte kan eller ska förbises.

Resultatet är ett antal verktyg och även råd för att kunna välja ut vilka produkter och eller tjänster där intjäningslogik, "the earning logic", föreligger med tanke på utveckling nu och för framtiden.

Språk: Engelska   Nyckelord: ICS, Data, Digitalisering, DevOps

---

MASTER´S THESIS

Author:	Andreas Bäck
Degree programme and place:	Technology Based Management, Vaasa
Supervisors:	Stefan Granqvist, Novia UAS
	Leif Strandberg, Wärtsilä

Title: Data driven decision making in Industrial Control System

---

Date 12.12.2019   Number of pages 58

Summary

This thesis is about benefits and values with data driven development in automation systems that are inside 4 - Stroke engines developed by Wärtsilä Finland Oy.

My main purpose is to both describe, explain and *to guide*, with a normative approach, how data can be used in order to make more analytical decisions for Industrial control systems. My secondary purpose is to both create and present tools and models for how to work with data in order to achieve the goals.

The thesis was carried out through review of literature and research on how to transform to a more data driven business. The research method can be categorized as an actor - based study method.

The result of this thesis are several tools helping Wärtsilä in identifying which products and services are most beneficial to be developed towards an even more data driven business linked to the company´s short term and long - term strategy.

Language: English   Keywords: ICS, Data, Digitalisation, DevOps

---

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

1 INTRODUCTION

This Master's thesis is part of the master's degree Program in Technology Based Management at Novia University of Applied Sciences. It strives to demystify the "Big Data" and "Digitalization" hype and see if it can be used in the area of Industrial Control Systems and Operational Technology.

To know if something is usable one needs to know what one should use this new technology for and whether there is money or resources to be saved and value added, both in a technology and a business context. If data can't be used for some reason, there is often only gut feelings and tacit information in use.

Let us make/take an analogue to the following statement:

Highest Payed Person Opinion "HiPPOs are anti – data - driven. They make decisions based on their experience, preconceived notions, and their gut, *without regard to the data available*" (Andersson, 2015, p. 214).

Figure 1. Hippos are among the most dangerous animals in Africa. Conference rooms too. (**Rosenberg, 2014**)

1.1 THE PROBLEM AREA

Energy production belongs to Common Infrastructure and there is a basic assumption that it always works in today's society. World is changing and we talk about speedy development. Those not following the trends or being incapable to keep up, have no future. Here follows a description of *four drivers on a macro level and four drivers on a micro level* forming a platform for thinking and for understanding of the problem area. I start with the first macro driver, *globalization.*

Today it is not crucial where you are born or where you live. We can study from almost anywhere, we can arrange meetings and conferences regardless of location – new phenomena's like "*digital nomad"* are not uncommon any more. Even if you can work from anywhere the world´s GDP production is moving towards a concentration into 600 cities. (Institute, 2011) Companies relocate or flag out constantly trying to be even more competitive. Decision makers going for the best strategies are confronted daily with the need for good data in order to make right decisions.

In a Wärtsilä context we are taking a path against autonomous ships in Smart Marine Ecosystem (Wärtsilä , 2019). When it is up to the level of autonomy there are several maturity levels existing and being defined in various industrial domains. Initially lower level of autonomous automations will be about diagnostics for preventive maintenance, predictive analytics towards measurable excellent operation and safety for human lives.

The other side of the coin is that in many parts of our world labor is cheap, making automation rollout initiatives slow (my own reflection Andreas Bäck 2019).

Globalization macro trend drives technologies in directions and actions as follows:

- Produce higher quality products (ISO9001, CMMI)!
- Do things faster and right things first (Agile, Rapid, DevOps, ROI, EBITDA)!
- Awareness of market reactions even faster (Flatter organizations, Big - Data, Speedy Development, Digital Twins, Agile, DevOps, DataOps, Business Agility)!
- Less practical competence level drives big data solutions to represent a new level of interactive feedback and insight (Industrial Operations, Dynamic Pricing, Remote Operations)!
- OPEX preferred over CAPEX (operational expenses versus capital expenses)!
  (collected information from literature, standards and internet, Andreas Bäck 2018 - 2019)

Second recognizable macro driver related to the problem area is t*echnology*. *Interne*t is all over the planet, intelligent devices are built into almost everything as technology prices have been reduced opening up possibilities and solutions in a new way. We see a logarithmical increase of connected IoT devices (Statista, 2017).

New materials and additive manufacturing and knowledge allow us to do things that weren't economically doable in the past. For example - cylinder pressure sensors on an engine are doable today, and this opens up ways to operate an engine that in the past only existed as a dream. Rudolph Diesel was not "there" at that time, meaning the tools were not there for implementing his excellent ideas! Part of the reason that Wärtsilä 31 engine is in Guinness world record is due to advanced sensors and automation (Wärtsilä, 2017).

The technology macro trend drives technologies in directions and actions as follows:

- Everything is connected and vulnerable (ISO27K, NERC-CIP, Grid-code, IoT)!
- Technology life cycles are shorter than ever (Obsolescence Management)!
- Complexity in business chains (Owners, Yards, Designers, Brokers, Licensees, Operators, Customer Centricity) enabling even new business models perhaps towards totally intermediate platforms (compare Uber, Alibaba, android platforms) (Bruzelius & Skärvad, 2017).

Third macro driver - *climate change* - can also be seen as a driver affecting my study´s problem area. Global warming combined with fossil changing fuel prices have led to new research in renewable technologies. Climate macro trend drives technologies in directions and actions as follows:

- Grid codes in countries change due to renewables!
- In Finland Fingrid has created an open Datahub - a centralized information exchange service for electricity trade!
- New businesses are born from political decisions (Circular Economy (CE)/Blue Economy (BE), The As a Service Concept, Carbon Bank)!
- Stricter regulations on both land and on sea (IMO Tier III Directives)
- Energy price can change from high to low in a day, which means the profits made one hour may turn into loss the next (Pooling, Big-Data)
  (collected from literature, standards, internet and (Günther, 2017)

The *fourth* macro driver is about *fear* in society. In handling the balances of power, we can identify new ways of handling for example protectionism of a country's business actions. Big companies and structures making big money do all they can to remain in the business with and via lobbying. The borders are open, IoT is global and the second machine age (Brynjolfsson, et al., 2014, p. 21) is in starting blocks. Fear macro trend drives technologies in directions and actions as follows:

- Manage Cyber Security (IEC62444, NERC-CIP, ISO27K)!
- Develop new business models, for example ownership is taking new shapes with all kind of As A Service "*AAS"! (see also page 3)!
- Manage the long - time organizational knowledge and try to reduce and identify the tacit information. (Obsolescence Management IEC EN 62402:2007)!
- Develop new cultures in companies like Agile, DevOps, Lean, Cross functional teams and DataOps that supports business agility!
- Centralization of companies around knowledge hubs towards specific concentration but also towards an even more focus on intermediate platforming in entrepreneurship (collected from literature, standards, internet, Andreas Bäck 2019) and (Bruzelius & Skärvad, 2017).

Now turning to the micro drivers. I have chosen to drill down from the macro - level cause all the drivers are connected to each other both seen from a system theoretical point of view and a parametrical point of view (Arbnor & Bjerke, 1994). I present the micro drivers starting with cyber security.

Cyber Security (CS) starts to be a mandatory requirement in product used for CI (Common Infrastructure). The CS here will be focused on OT (*Operational Technology*) leaving the IT (*Informational Technology*) for other studies.

There exists a magnitude of standards like ISO27K and NIST SP 800. The ones that are used in Wärtsilä are NERC - CIP and IEC EN 62443. OT side of CS has much more in common with Functional Safety (FS) than IT, meaning in OT the primary concern is to keep "the light" on. IT is more about ensuring that the private data is kept private, e.g. passwords are correct and ensures that the right persons do what they are allowed to do. A General Data Protection Regulation (GDPR) legislation became active 25 May 2018 from EU. Majority of GDPR is for the IT part of organization, but care must be taken when one collects information from OT side if it's possible to link to a person. On the other hand, it is a requirement in CS standards to know who have done what.

Cyber security in practice means the following:

- Operators must be able to ensure that vulnerabilities are mitigated during the lifetime of system
- Risk assessment of threats must be in order. (Risks are Level 3 in CMMI-SVC 1.3 and maturity levels are aligned with IEC 62443-2-4:2017). ISA and IEC have aligned a standard, **see figure 2 on page 6!**
- Quality must be on same maturity level defined for what one intends to have for cyber security maturity level.
- Fulfill mandatory requirements from classes and states/governments in order to continue using products, systems and services avoiding penalty fees.
- Asset management so that one knows what is installed and where.
- A smooth working organization that can deliver security patches in a fast and timely manner.

Figure 2. ISA 62443 organizational structure (Knapp & Langill, 2015, p. 393)

The next micro driver is *lifetime alias* obsolescence management. IEC 62402:2007 is a standard that exists in order to help with defining the lifecycle of a product. With faster iterations on technology combined with an industry that expects 30 or even 40 of years lifetime for an engine, the standard sets a clear focus on having organizations that is competent to plan, monitor and build future proof solutions. IEC 62402:2007 standard is primary focusing on hardware components, but it does not say that the principles cannot be used on/in other areas.

Depending on who you talk with, domain language may differ, and you may hear words like *life cycle management*, *future proofing*, *portfolio management*, *product management, obsolescence management and Product Lifecycle Management* (own reflection Andreas Bäck 2019).

Here follow *challenges* related to obsolescence management:

- How to be able to produce automation equipment resources effective for at least 20 years when the longest time you can get electronic components is approximately 10 years?
- How to ensure building environments that are still working for 20 years or "what if" "you need to be able to produce updates and patches for the product?
- How to ensure organization and people still maintain competence for a product that is not done daily anymore? Here the company also must ensure that you have enough "bus factor" (Fitzpatrick & Collins-Sussman, 2012, p. 7) and not only basic knowledge (own comment Andreas Bäck 2019).

- Ensure with annual verification that the organization is still able to produce the product and the services.
- Being agile and assemble specialized groups solving new business ideas rapidly and at the same time making sure there is a back office taking care of all the loose ends required for quality in the long run.
- Planning for upgrading/replacing of products with newer ones in a planned manner.
- Develop a forward proof architecture that allows continuous deployment and innovation.

The third micro driver is *digitalization*. Here I mean Big Data, IoT, Analytics and DevOps needed for going from a *technology driven organization* and more to a *data driven organization* that delivers business values with minimal waste and maximal profit. **There is a lot of data around in the companies but is the data used effectively**? An important question is "Do You have control of Your Data"? How do we put important and useful data into - "real use"? (own reflection Andreas Bäck, 2000 - 2019).

The Big Data that later evolved into the analytics hype (compare Gartner 2010 - 2019) is a clear sign that at least some companies can make things smarter by using the data in their systems to make better decisions and not take the Highest Paid Person Opinion (HiPO) path. If decisions are based on only gut feelings you are using the "fast brain" in a person that has been lucky (Wiseman, 2004), and he /she will make decisions based on previous decisions if no data exists. If we can acquire good data for the "slow brain", we can make improved analytical decisions (Kahneman, 2013).

In automation there exists a lot of uses for digitalization that can assist and be **put into use**. Automation also has the benefit that the quality of data can quite simply fulfill the data quality issues mentioned in (Redman, 2008, p. 41), see figure 3 on page 8.

Issue	Benchmark
People can't find the data they need	Knowledge workers spend 30% of their time searching data they need, unsuccessfully half the time.
Incorrect data	10–25% of data record contains inaccuracies
Poor data definition	Data frequently misinterpreted
	Can't connect data from different departments
Data privacy/data security	All data subject to loss
	Risk of identity theft
Data inconsistency across sources	The norm when there are multiple databases
Too much data	Half of all data never used for anything
	Uncontrolled redundancy
Organizational confusion	Can't answer basic questions such as: 1. How much created each day? 2. Which are most important?

Figure 3. Seven common data quality issues and benchmarks for each (Redman, 2008, p. 41).

Useful data linked to automation & digitalization issues are:

- Who commissioned or visited this product, GDPR papers in shape!
- Full trace of production data (Industry 4.0 and ISA95)!
- Standards interfaces to make integration faster and possible!
- How many instances of this vulnerability exist in the field? Important question!
- What functionality is most/least used by users?
- Be aware of problems and mitigate before escalated!
- Help stakeholders and customers earn more money!
- Is usage pattern of product changing?
- Improve overall quality, by knowing all potholes and doing right - quality will follow!
- Reduction of false positives, in other words clean the data for AI already on the field!
- **Understand how customers and their customers use our/your products/services! Understand the chain!**

IT industry has for quite some years already been doing *digitalization* and that means that software that you traditionally in the past bought on a CD and installed on a server also had a specialist locally that understood this software and modified it to suit your organization. Software as a Service (SaaS) and other "as a Service concepts" are increasing in all companies today. This enables all sizes of companies to use same software and that enables a faster growth of the company.

SaaS benefits and constraints are as follows:

1. Pay per month! / License instead of initial fees on both software, hardware and competence, meaning the risk is transferred from the end customer to the licensee.
2. You can change software whenever you like as long you have made sure data that is still needed can be transferred.
3. You don't need to find local competences for a software as the company making the software is hosting it.
4. Interfaces need to be standardized so different vendors can exchange data.
5. Challenges with SaaS comes from when having long (over 5 years) *lifetimes* on products, as there exists very few actors that want to provide backward comp ability and even promise longer contracts.
6. A change in the skillset from technical to more making contracts and a holistic architect view of the services one use and their data flow.
7. The licensee will need to understand how to run things optimally and improve the efficiency during the whole lifecycle instead under warranty time based on data collected.
8. Speed up the learning and innovation from data, the faster one can get data back from field to R&D one can faster innovate and react to changes.

There is also a possibility to create new businesses as the company has collected data and that data can also be valuable for other players, and even new players, potential customers, that have not existed before. Out there exists amounts of already collected data and if there is lacking data that data can today be quite easily acquired for example via Fingrid's open data initiative.

1. Fingrid open their data (Fingrid, 2017) and this will create new business trying to sell ways to reduce your electricity bill by helping you choose when to consume electricity.
2. If ships would know well in advance about the que time in the harbor and you create a market out if this, ships can then plan ahead and depending on what they delivers to wait or pay someone else to wait and the ques will optimal. (Michaelides, et al., 2019)

The last micro driver linked to my study area is *Safety*. IEC EN 61508 is the main standard and there are several industry specific adaptions like Road Vehicles IEC 26262, Medical device software C 62304 that many products must follow in order to make sure that people are not injured or die linked to usage of your products.

Failure Mode and Effects Analysis (FMEA) tries to find possible faults, and are both used for defining the demands, and also serves as input for several architecture choices, like in cyber security network segmentation the safety zone needs to be of its own, and sensors are really not allowed to share data/signals between safety part and the controller part. (Smith & Simpson, 2011, p. 48)

Shortly described functional safety is a quality standard with Safety Integrity Levels (SIL) 1 to 5 where 5 is most complex and demanding. "There is a considerable body of opinion that SIL4 should be avoided and additional levels of protection should be preferred". (Smith & Simpson, 2011, p. 10). Based on how big a risk there is for people together with how well in advance one has been thinking of the architecture gives *minimum safety integrity level*. Safety is a very regulated domain and as such has a very conservative view on how development shall take place. As of writing the recommended development model is the V - model (Smith & Simpson, 2011, p. 69) see figure 4 and there are still only some researchers (own reflection) studying if it could be done agile. This means that we are many years away from using data in the same way as in other domains like within Machine Learning (ML) and Deep Learning (DL).
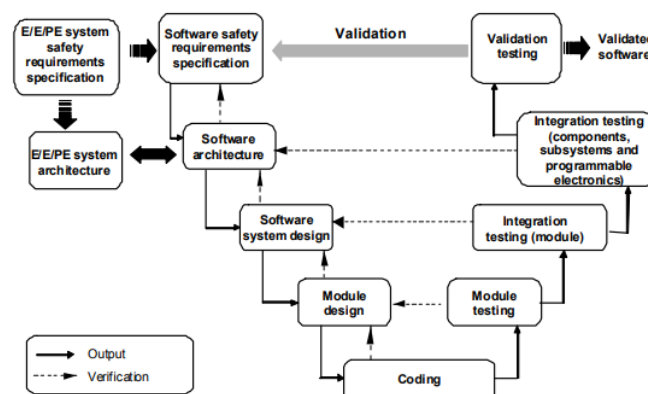

Figure 4. Software systematic capability and the development lifecycle, the V-model **(IEC, 2010, p. 17)**

Important when it is up to functional safety, I underline the following:

- What function is the weakest link in the chain? Find it and define it!
- Recorded statistics from field and production about subcomponents, describe, explain and analyze for robust decision making!
- Traceability from all events that made up the product and ensure that they are in synchronization with the defined quality process!
- Coverage of tested functionality!
- Data for investigation if something goes wrong!

In figure 5 I have tried to show Macro and Micro drivers and the relationship between them. There has also been an attempt in this picture to have the related macro and micro drives as close to each other forming bridges as follows:
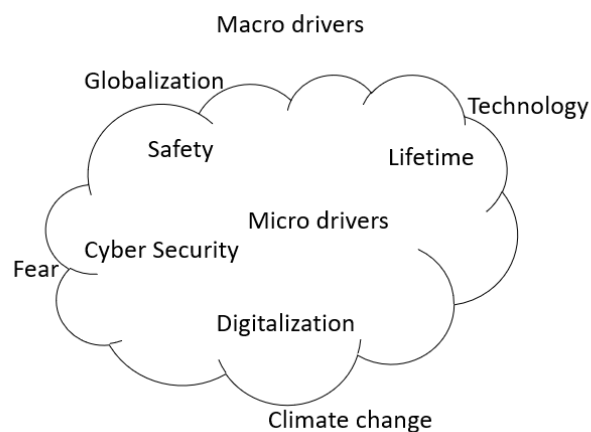

Figure 5. Macro & micro drivers Andreas Bäck 2019

1.2 POSITIONING THE RESEARCH AND THE RESEARCHER

I am working as a UNIC SAFe Product Owner (PO) in transformation, current in Product Increment (PI) 4. Each increment is 10 weeks for the automation system UNIC (Wartsila, 2017).

I started my career in the company 13 years ago as a design engineer developing engine software for Spark ignited Gas (SG) 4 - Stroke engines. Before that I worked 5 years as a PMI (Project Management Institute, 2019) certified project manager for the Wärtsilä Modular Application Platform (WMAP) that is the operating system for UNIC.

My Bachelor´s thesis in year 2000 was about version control and Capability Maturity Model (CMM) that is also used in this Master´s thesis as a fundament for preunderstanding.

This my Master´s thesis can be categorized as *an action based study* (Runeson & Höst, 2008, p. 136) carried out by a person interested in developing and contributing to the firms' quality and overall growth for success. The study is carried out for Wärtsilä Marine Business specifically for the Department of Automation and Control.

1.3 PURPOSE OF THE STUDY

My main purpose of the study is to both describe, explain and present *advices where and how improved data quality* can be used for better analytical decisions in Industrial Control Systems (ICS) in Wärtsilä in order to improve the company´s short- and long - term profitable business.

My secondary purpose is to, if possible, make a first attempt to create a model of way of working with data. The attempt can result in a form *of a model, a map* or set of building stones for a new improved system linked to the department strategy of acting and the company strive for an even more measurable value orientation.

1.4 LIMITATIONS

I limit this thesis to where 4 - Stroke engine automation systems are part of the solution in the Marine Business segment. The company I work for have all its business internationally and several of the drivers are global standards and mandatory requirements from stakeholders that need to be fulfilled in order to be part of the challenging markets.

I will limit my descriptions, explanations and advices into data that one can make a user story for. Background to this limitation is both the General Data Protection Regulation (GDPR) and storage can be quite expensive when considering backups. Having an agile mindset, one shall focus on user stories in order to avoid gold plating.

My focus in this Master´s thesis is on the big company´s processes and I am not trying to orientate into micro companies or small and medium sized (SMEs) companies. That orientation needs another thesis.

1.5 MY THESIS

It is now right time to carry out my Master´s thesis because I and many a researcher are in a turning point into a technological revolution and transforming into an *Age of Software & Digital* (Kersten, 2018, p. xiv). The facts that touch many business companies today is addressed in the following figure 6 as follows:
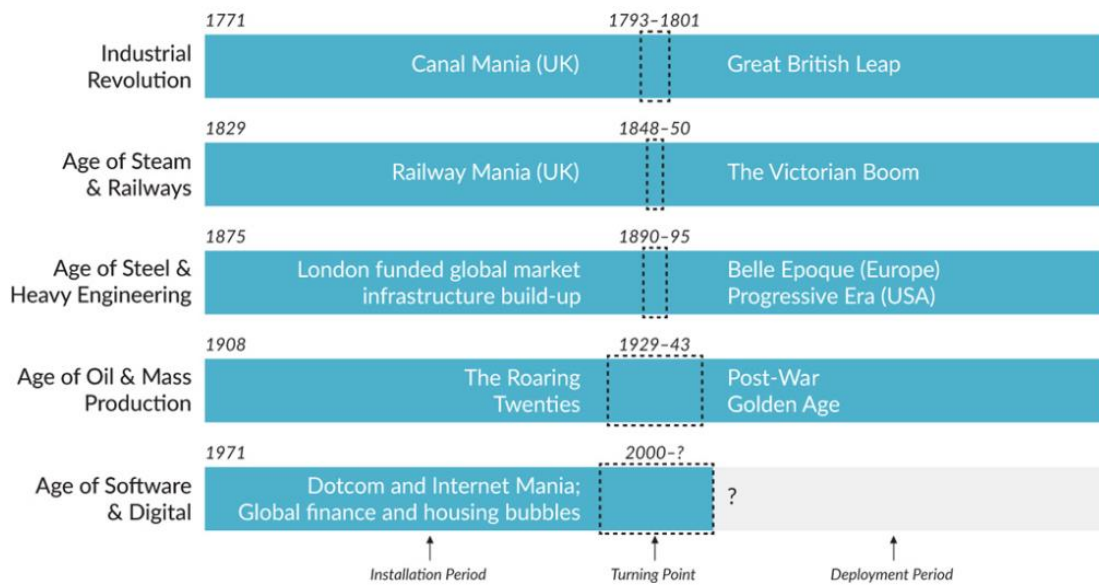

Figure 6. Technological Revolution and the Age of Software **(Kersten, 2018, p. xiv)**

Wärtsilä, as a world leading company going for digitalization with help of operational excellence and value orientation and the need for continuous development is on the daily agenda and tightly underlined in the company´s future strategy named Smart Marine Vision. (Wärtsilä, 2019).

1.6 TERMINOLOGY, ABBREVIATIONS AND ACRONYMS

In this section I explain terminology as follows:

- **AMR**: Automatic Meter Readings is technology used for collecting in usage data from consumers about electrical/water and similar usages.
- **ALM**: Application Lifecycle Management is the products lifecycle management.
- **IOT:** Internet of Things
- **CMM:** Capability Maturity Model is quality standard developed by Carnegie Mellon University for US Department of Defense.
- **CMMI:** Capability Maturity Model Integrated is continuation of CMM when it integrated 2002 several domain standards into fewer ones. The 5 step maturity levels used in many places have its origin here. [1 Initial, 2 Managed, 3 Defined, 4 Quantitatively Managed, 5 Optimizing]
- **ICS:** Industrial Control Systems is for software and hardware systems used in industry. Have higher grade of demand to work 24/7 than consumer grade electronics.
- **ICSs:** Industrial Control Systems safety is word used in ICS when talking about CS. Difference to CS is here, we focus on keeping the systems running when under attack compared to CS where we shut down.
- **CS:** Cyber Security or IT security is the protection of information technologies
- **IT:** Information technology
- **OT:** Operational Technology, also called ICSs before.
- **NERC:** North American Electric Reliability Corporation is the northern part US electrical grid company that creates all standards before they are federal. Counterpart to this in Finland is FINGRID.
- **FERC:** Federal Energy Regulatory Commission, Whole US grid.
- **DASLO:** Acronym from Develop, Active Supported, Limited, Obsolete obsolescence lifecycle
- **CIP:** Common Infrastructure Protection is a standard by NERC/FERC for ensuring critical electrical network have resilience against malicious events.
- **Functional Safety:** Freedom from unacceptable risk of physical injury or of damage to the health of people, either directly, or indirectly as a result of damage to property or to the environment. (IEC 61508, 2010)

- **SIL:** Safety Integrity Levels defines how much effort is needed on safety for a function.

- **HiPPO:** Highest Payed Person Opinion, Hippos are among the most dangerous animals in Africa. Conference rooms too. (Rosenberg, 2014)

- **Digitalization:** is the use of digital technologies to change a business model and provide new revenue and value-producing opportunities; it is the process of moving to a digital business. (Gartner, 2017)

- **GDPR:** EU's General Data Protection Regulation rules about person privacy data active from 25 May 2018 (European Union, 2017) .

- **SaaS**: Software as a Service

- **SAFe** Scaled Agile Framework

- **DataOps:** an automated, process-oriented methodology, used by analytic and data teams, to improve the quality and reduce the cycle time of data analytics.

- **DevOps:** Development and Operations.

- **Business Agility** ability of a business system to rapidly respond to change

- **3PLM** The Three Layer Product Model

- **CDO** Chief Data Officer

- **Age of Software:** the current technological revolution that begun in 1971, and is marked by advances in microprocessors, telecommunication, the internet, and software. (Kersten, 2018, p. 213)

- **ITIL:** Information Technology Infrastructure Library

- **Bus Factor:** The number of people that need to get hit by a bus before your project is completely doomed. (Fitzpatrick & Collins-Sussman, 2012, p. 7)

- **ISO:** International Standards Organization

- **PMBOK**: Project Management Body of Knowledge by PMI

- **CALMR**: Culture, Automation, Lean, Flow, Measurement and Recovery

- **PI**: SAFe Program Increment.

- **PO**: SAFe Product Owner.

- **B2D**: Business to Digitalization (Peter Fader 2012 and Predictive Analytics World, London 2015 – 2018, Sven Krohne)

- **CVE:** Common Vulnerabilities and Exposures (NIST, 2019)

- **CPE**: Common Platform Enumeration (NIST, 2019)

- **ISA**: Industry Standard Architecture

1.7 DISPOSITION

Here is the disposition of the thesis as follows:

Chapter two presents the theoretical frame of the thesis for both my preunderstanding and insight for further descriptions, explanations and advices.

Chapter three is about the method used in this thesis.

Chapter four presents tools that help decision makers in understanding which products and services to go forward with in order to become a data driven organization striving for the vision to introduce a B2D (Business to Digitalization) paradigm compare (Fader, 2012).

Chapter five presents the conclusions of this thesis and suggestions for further research.

2 THEORY BUILDING

This chapter is about theory building linked to development processes in an industrial context. There is a lot of books, research reports and journals linked to new technologies such as digitalization, systems thinking and new business creating models. Additionally there are training courses arranged by professionals and of course in a multinational company in - house training continuously. The study area and my thesis problem area can be seen as a multidisciplinary canvas where many sciences are linked together meaning that data science, analytics, digitalization, economics and even psychology are struggling side by side.

If a company´s goal is to deliver quality and be profitable and at same time having business agility, then question is "what data and practical actions even down to transactional level is needed to make right decisions?"

We must observe that when I am presenting theories and facts I sometimes here comment and build bridges to what happens in Wärtsilä today. I am aware that theory building, and field observations sometimes cannot be mixed into each other but here I argue that theoretical findings and experiences can be matched cause the combination "teach us to reflect and learn"! The nature of this action-based study and the fact of interdisciplinarity in a scientifically sense of view defends my standpoint.

Section 2.1 is about benchmarking and gap analysis to select out what successful companies are doing and how. In section 2.2 macro and micro drivers are described as important building stones for budgeting and resource usage linked to development processes. Section 2.3. presents the life cycle management process. In section 2.4. the stair heaven model is described. Section 2.5. is about value stream mapping. 2.6. is about holistic development framework and in section 2.7. follows a short summary of the chapter.

2.1 BENCHMARKING & GAP ANALYSIS

The first thing to understand is in what domain your product and solutions are in. For example in Industrial Control Systems we seldom can go to the extreme and skip requirements on the products and find out what customers wants from A/B testing. Benchmarking and gap analysis provide basic insight how a company sees and uses lessons learned from another company and often it is worthwhile to benchmark a successful actor. Gaps are identified and at the same time it possible to avoid mistakes or to repeat mistakes already done.

The model by Ståhl & Mårtensson 2018, in figure 7, is a starting point when critically gripping crucial development actions in the industrial context. This model summarizes the difference in the domains, and it would be a considered alternative to use in Wärtsilä when deciding who to benchmark. Another import thing is that this model can be used when purchasing extern consultants for development tasks. The model also serves as a tool when bringing in new people that come from other domain areas. The model is described in both table 1 and in figure 7 on page 20.

My comments:

"Many of the products/solutions in the UNIC portfolio that I am looking at in this thesis share many attributes with Automotive Industry Segment and archetype Alice referring to table 1 on page 20. Some modules we have for Functional Safety is a little bit higher on Safety, and data collection is not allowed from Classification Societies and Cyber security regulation in the near future. Wärtsilä has other tools like UNITool/ WecsplorerUT, also called maintenance and monitoring tools and these two don't have as much formal requirements but can collect more data and can be developed more like the industry segments *social media* and *computer games*, see table 1 on page 20. GDPR requirements are to be fulfilled." (Andreas Bäck 2019).

Table 1. Archetype summary **(Ståhl & Mårtensson, 2018, p. 44)**

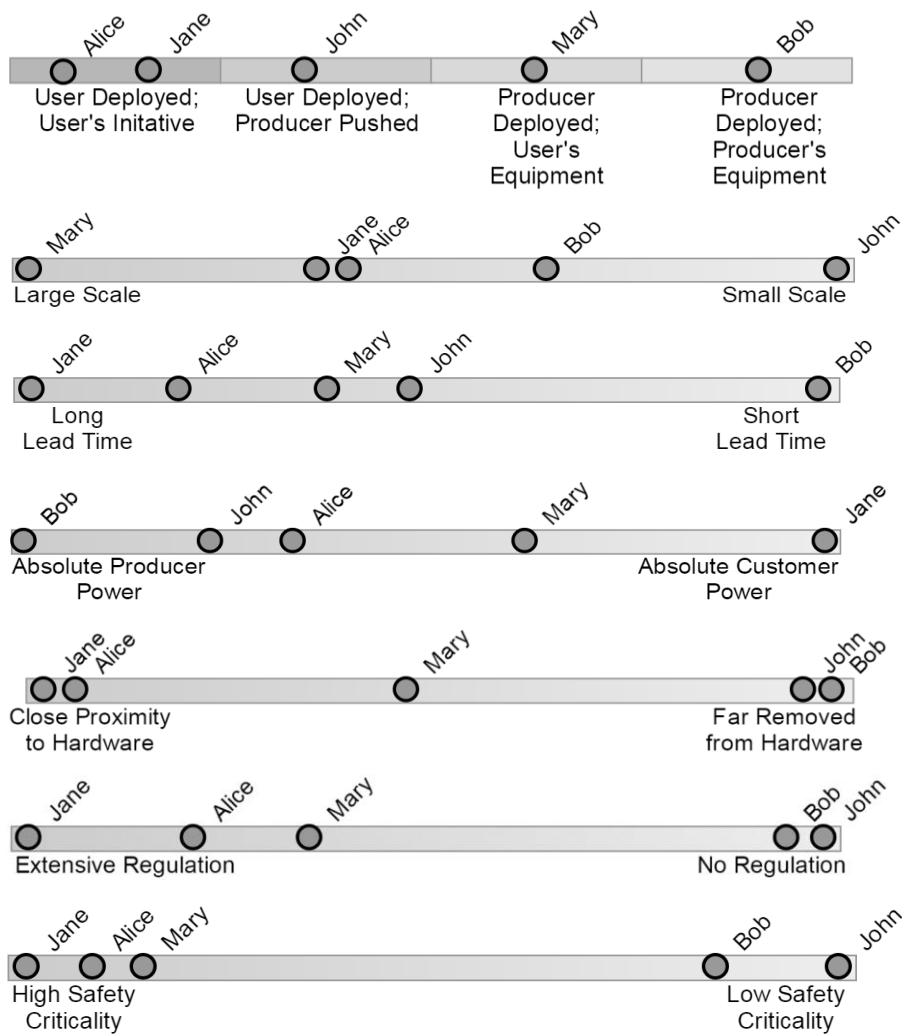| Archetype	Industry Segment
Jane	Defense
Bob	Social media
Mary	Networks
Alice	Automotive
John	Computer games

Figure 7. Collection of pictures from **(Ståhl & Mårtensson, 2018, pp. 33-100)**

2.2 QUALITY DEFINITIONS

In this section I am referring to important building stones from literature forming the foundation for quality development and statements in an industrial context. Quality and six important definitions starting with the first definition are as follows:

ISO 9001: "Degree to which a set of inherent characteristics fulfills requirements" (ISO/IEC 9001, 2015)

Six Sigma: "Number of defects per million opportunities." (Anon., 2017)

CMMI-SVC: The quality of a system or product is highly influenced by the process used to develop and maintain it". (Software Engineering Institue, 2010, p. 5)

Agile Manifesto Principle #1: "Our highest priority is to satisfy the customer through early and continuous delivery of valuable software". (Beck, et al., 2001)

DataOps Principle #1: "Continually satisfy your customer". (DataKitchen, 2019)

Wärtsilä: "Our vision is to be our customers' most valued business partner'. (Wärtsilä, 2017).

My comments: Among the many definitions existing for quality I therefore choose to use the most fitted one for me and my research context. I here choose to go for what Wärtsilä promises its owners: "Our vision is to be our customers' most valued business partner " (Wärtsilä, 2017).

One definition of quality is perhaps not enough in a big company with many divisions and multitude of suppliers. Six Sigma is used by Wärtsilä especially in the supplier's interfaces for products and components. If we talk about internal development and suppliers delivering services ISO 9001 is more used and this is also the one that classification societies follow. CMMI and its sub variants have for a long time been the guiding start in both standards like Security standards like (IEC 62443) and for industry segments.

2.3 PRODUCT LIFE CYCLE (PLC)

Wärtsilä is using a product lifecycle model created by our self-named DASLO loosely based on how ABB Drives have defined it (ABB Drives services, 2016). When the model was developed standards like obsolescence management (IEC 62402:2007) was influenced to DASLO model and that is also the reason that it's very hardware centric. There exists in Product lifecycle literature several models, but according to my findings it only describes a subset of these life cycle phases. This lifecycle is primary developed to support how automation hardware modules are planned. If one uses this same model for software one has a traditional lifecycle thinking and that is a start, but it requires some more work in order to support e.g. As a Service (own comment Andreas Bäck 2019).

Table 2. Mapping of Wärtsilä' s DASLO and ABB Drives Andreas Bäck 2019

Wärtsilä	**D**evelop	**A**ctive	**S**upported	**L**imited	**O**bsolete
ABB Drives		Active	Classic	Limited	Obsolete
IEC 62402:2007		Introduction	Decline	Phase out	

Here follows what DASLO stands for as follows:

Develop - in this stage a product is developed**,** depending on the component this can take quite some time if it's not a component found on the shelf and have compatible replacements from other suppliers. In our (Wärtsilä) case this can be everything from 0 – 4 years for automation.

Active is the time the product is manufactured and available from factories. Normally we calculate this to be 10 years for Automation and 20+ years for engines.

Supported is no longer the default from factory, Customers get support and spare parts. Normally calculate this to be 10 years for Automation and 20+ years for engines! (own norm!)

Limited spare parts can no longer be guaranteed, and customers are informed and can plan what they should do next.

Obsolete - we no longer keep competence, documentation and spare parts.

Digitalization requires Project to **Product oriented management** focus:

As we already saw from previous, the planning of the whole lifetime of the product becomes more important and that leads automatically to a shift from Project Management to Product Management focus. To understand what this means in practice we can look at figure 8 onto the different areas and what impact is for you:

	Project-Oriented Management	Product-Oriented Management
Budgeting	Funding of milestones, pre-defined at project scoping. New budget requires creation of a new project.	Funding of product value streams based on business results. New budget allocation based on demand. Incentive to deliver incremental results.
Time Frames	Term of the project (e.g., one year). Defined end date. Not focused on the maintenance/health after the project ends.	Life cycle of the product (multiple years), includes ongoing health/maintenance activities through end of life.
Success	Cost center approach. Measured to being on time and on budget. Capitalization of development results in large projects. Business incentivised to ask for everything they might need up front.	Profit center approach. Measured in business objectives and outcomes met (e.g., revenue). Focus on incremental value delivery and regular checkpoint.
Risk	Delivery risks, such as product/market fit, is maximized by forcing all learning, specification, and strategic decision making to occur up front.	Risk is spread across the time frame and iterations of the project. This creates option value, such as terminating the project if delivery assumptions were incorrect or pivoting if strategic opportunities arise.
Teams	Bring people to the work: allocated up front, people often span multiple projects, frequent churn and re-assignment.	Bring work to the people: stable, incrementally adjusted, cross-functional teams assigned to one value stream.
Prioritization	PPM and project plan driven. Focus on requirements delivery. Projects drive waterfall orientation.	Roadmap and hypothesis testing driven. Focus on feature and business value delivery. Products drive Agile orientation.
Visibility	IT is a black box. PMOs create complex mapping and obscurity.	Direct mapping to business outcomes, enabling transparency.

Figure 8. Project-Oriented Management Vs. Product-Oriented Management (Kersten, 2018, p. 54)

In traditional product development that value of the product starts to degrade as soon as it has been delivered to the customer, whilst in a software age with "continuous value streams" the value of the product is even improved with software updates until the hardware can't support it any more. Inside each step in DASLO typically normal project driven projects have been the norm "today", but all predictions show that best way to deliver in the "Tomorrow" is to deploy all the time as figure 9 implies:
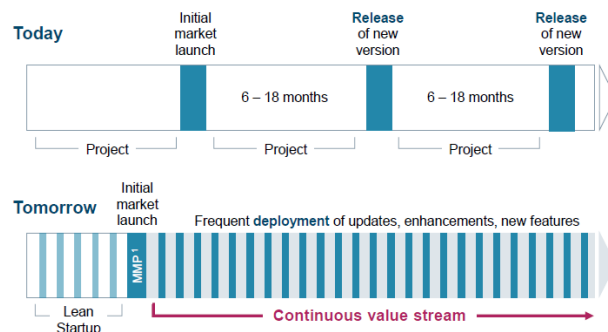

Figure 9. From project to continuous value stream **(AG, 2019, p. 6)**

From a *lean budget* view looking at this *"The Three Product Layer Model"* (3PLM) in figure 10 R & D resources used are usually described as *commoditized* part of development usually taking up to 80 - 90% of the R & D budget. (Bosch, 2018, p. 9). This makes it very hard to have any big profit margins as the customer assumes these should work. As a rule of thumb 10 % should be innovation and 50/50 division between differentiation and commodity (Bosch, 2018, p. 71).
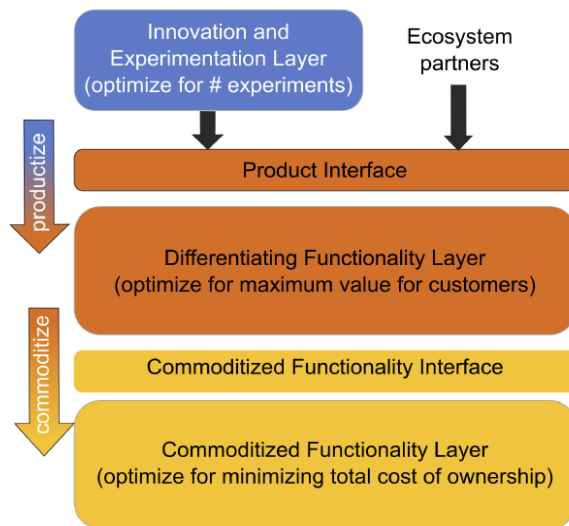

Figure 10. The Three Layer Product Model **(Bosch, 2018, p. 96)**

It's traditional budgeting that R & D budget is getting a percentage around 4 % of <u>EBIT</u> for investments. Investments in R & D normal demands that for "*every € spent has to result in 20€ of revenue for the company* " (Bosch, 2019, p. 110). This kind of budgeting becomes a challenge in the future as the amount of cost of software development is taking a bigger portion all the time. If one takes as example figure 11 from Car Industry software cost will be more than half of the product price due to that software will be updated continuously and enable new business models and keep the product "fresh".
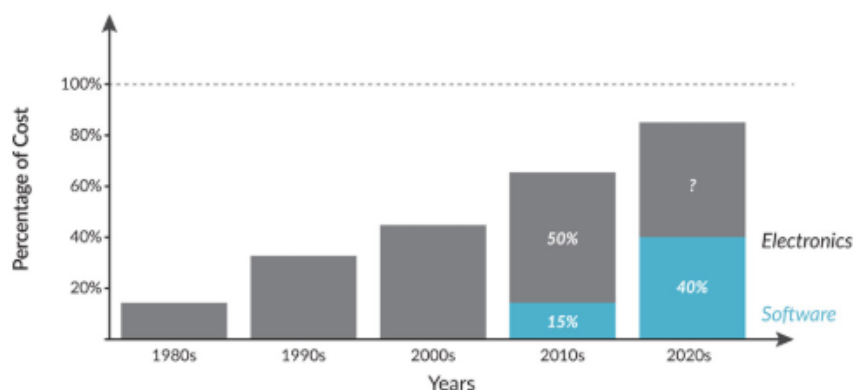

Figure 11. Software as Approximate Proportion of a Car Cost **(Kersten, 2018, p. 14)**

From a *DataOps* view this are usually called "Data Defense" and are defined "Data Quality, security, privacy, governance, compliance – these are all critically important endeavors, but they are just enabling activities" that provide indirect value (Bergh, et al., 2019, p. 59). Looking at this from a Chief Data Officer (CDO) position these can very easily divert to much focus from "offensive data", and if one wants to keep CDO position longer than 2 years one needs to show profit and these defensive data usually have too long cycle time to show fast enough progress. (Bergh, et al., 2019, p. 60). So, we need to find the right mixture of offensive data alias direct in order to have short term cashflow in order to finance the longer time to achieve profit projects (own reflection Andreas Bäck 2019).

Business case is the key to why the company shall invest to "just" collect some data. The best place to start building the business case is to talk with the people handling the finances in the company and it's highly suggested they should also be the owner of this "change program" as they know what the real cost of things is.

Hard costs found from the book keeping:
- Current lifecycle costs.
- Current nonconformity cost.
- Competence costs.
- Number of Non-Standard Requests (NSR)
- Identification of cost for overlapping products and lifetime. (Skärvad & Olsson, 2017)

Soft costs / benefits that need to be estimated.
- Increased innovation speed.
- New business models like as a service
- New businesses from trading of data with partners and others.
- Keep customers more satisfied.
- Enabling a platform for the whole ecosystem aka industry to build around.
    (Summarized from theory chapter 2 Andreas Bäck 2019)

2.4 STAIRWAY TO HEAVEN

Speed or the "Analytics Cycle time" is of the biggest importance in a data driven organization as it the will give the pace into all innovation and lessons learned from Data and AI. Digitalization is described in this maturity model as "*(software, data & AI) is disrupting industry and society to an extent that we only have seen the early beginnings of."*

I will use a maturity model "Stairway to Heaven" (STH) by Bosh in his Book "Speed, Data, and Ecosystems" is. Primary use for this STH model is to find out where one is now and what next the step to take in the maturity model in all the areas of Speed, Data, AI and Ecosystems. As many other maturity models like in CMMI the idea is to walk the stairs in all the areas in same the level, the reasoning behind that is if you are very mature in Data and AI but speed is lagging behind in Speed then you can't make use of the learnings and get the benefit delivered to customer and start to learn again.

The figure 12 below is a more condensed version from the book (Bosch, 2017, pp. 8-10) with addition of color coding in order to clearly indicate what staircase one talks about in the Stairway to Heaven (STH).

Speed	Traditional	Agile	Continous integration	Continous Deployment	R&D as innovation
Data	Ad-Hoc	Collection	Automation	Data innovation	Evidedence based org.
Ecosystem	Internally Focused	Ad-Hoc	Tactical	Strategic Single	Strategic Multi
AI	Experimentation & Prototyping	Non-critical deployment of ML/DL components	Critical deployment of ML/DL components	Cascading deployment of ML/DL components	Autonomous ML/DL components

Figure 12. Stairway to Heaven **(Bosch, 2017, pp. 8-10)**

The idea with this STH model in figure 12 is that one identifies where one is today on all the 4 areas and that starts to make a plan where you would like to go. I will only go deeper into the AI part of this model in figure 12 above as that is the newest based om both theory and existing practices so far (own reflection).

2.4.1 ARTIFICIAL INTELLIGENCE

One of the most hyped areas today is Artificial Intelligence (AI) with the subdomains Machine Learning (ML) and Deep Learning. (DL). DL is facing a couple of challenges as follows:

Algorithm development is by looking at the data and having neural networks or algorithms performing the work that in the past was done by a software engineer. Looking at figure 14 on page 28 we can conclude that interest in algorithms is existing today and there start to exist all the development and tools for doing algorithm development. Using these tools today is merely a question of competence development. Applying ML for e.g. using TensorFlow on a toy car Donkey Car (DonkeyCar, 2019) is no bigger challenge.

Platform is a combination of several parts in order to get the data in and out to the AI. In figure 13 tries to depict what challenges exist to get an AI driven development and this approach I have picked out from the paper for the automotive industry "Case A Domain Automotive" and which the challenges are (Lwakatare, et al., 2019).
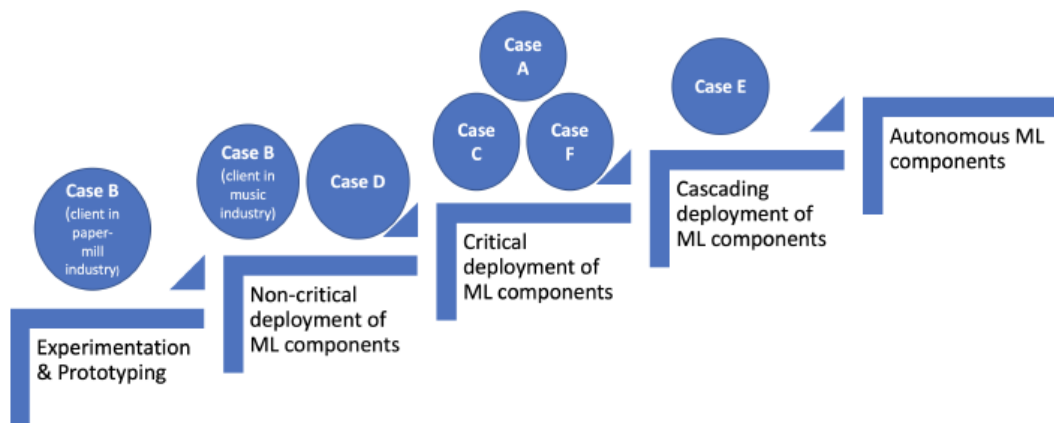

Figure 13. Evolution of use of ML in commercial software - intensive systems **(Lwakatare, et al., 2019)**

Case A Domain Automotive:

> *"There are no tool-chains you can download in an infrastructure with deep learning like this. And we realized after the mistakes and discussions with our new IT that they didn't really have the expertise to be able to deliver this to us. So we had to create new teams, which took the responsibility of creating both the infrastructure, but also the software tool-chain to be able to train deep learning networks within a reasonable amount of time".* (Lwakatare, et al., 2019)
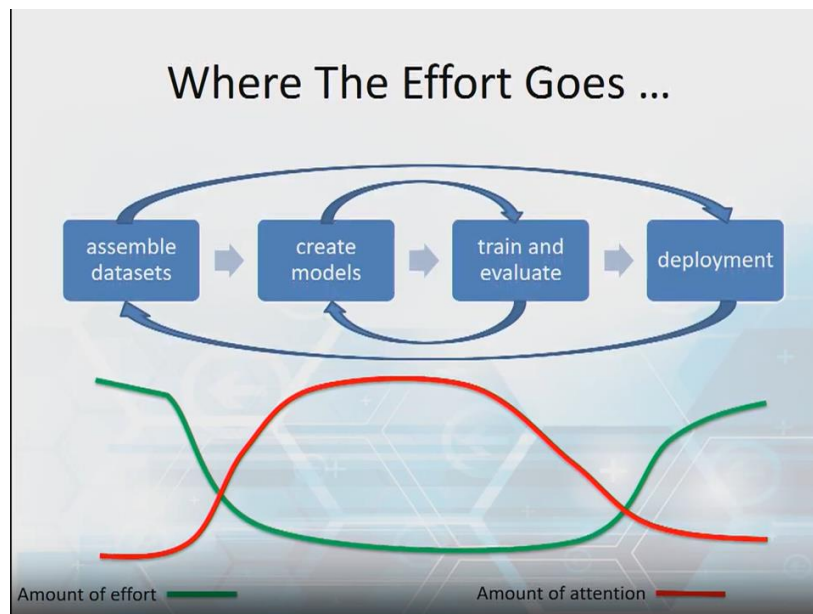

Figure 14. Where the effort goes (**Bosch, 2019, p. 16:47**)

As we see from the figure 14 ML and DL are getting a lot of attention and funding for creating new algorithms, but the actual effort to really get a working production ready infrastructure is **lacking** and it's first when we can have a closed loop see figure 15 from AI/Data driven decision making the process will start to pay back the investment.
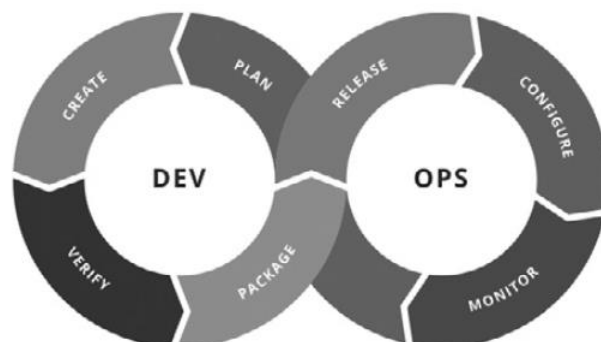

Figure 15. The DevOps lifecycle is often depicted as an infinite loop. (**Bergh, et al., 2019, p. 36**)

Collaboration between different functional domains is perhaps one of the biggest challenges with a Data driven (AI or Outcome) development. In figure 16 I explain that there is a bigger need for collaboration than before. IT department working with quality standard like Information Technology Infrastructure Library (ITIL) (AXELOS, 2019) and the automation department working with automation need to even more start to work together and be in tighter dialogue! In the past the role of the IT department was merely to provide standard IT equipment but that is no longer the truth: They are in a key role in providing value to the products in the organization. In the past it was not unusual that the R & D department had to have its own IT department as the IT department could or would not support the infrastructure and software that R & D needed. Today one needs to move in the cloud in order to use the resources efficient and smart. Figure 16 is also showing that in industrial systems we can't copy right of the concepts from original DevOps! We have to adopt the right side, see figure 16!
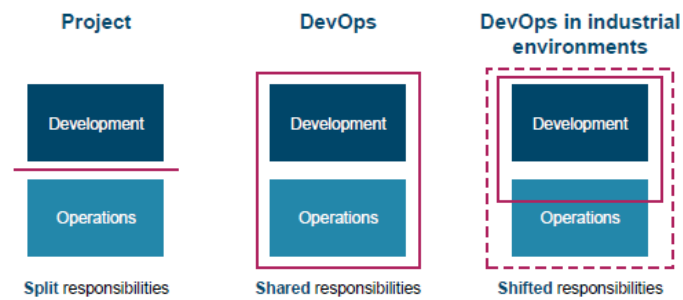

Figure 16. The DevOps organization **(AG, 2019)**

2.5 VALUE STREAM MAPPING

A Lean tool commonly used is value stream mapping to identify the cycle time, waste, flow of your product and how it will generate business value. From the SAFe framework definition is *"each value stream is a long-lived series of steps used to create value—from concept to the delivery of a tangible result for the customer"*.
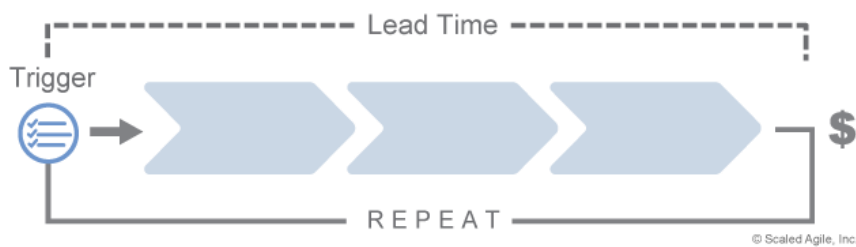

Figure 17. Anatomy of a Value Stream **(Scaled Agile, Inc., 2018)**

From this tool from a data point of view crucial questions are as follows;

- What is the waste found in the value stream?
- How and where is data stored from the stream?
- Where is time spent waiting?
- What is Lead time and cycle time?
- Is the data normalized and can it be shared between competence teams in the organization? (summary of the wastes defined in Lean regarding data, Andreas Bäck 2019).

The official lean wastes that forms the acronym DOWNTIME are also beneficial to take a look on while doing this value stream exercise that figure 18 shows:

Figure 18. 8 Wastes (goleansixsigma, 2019)

2.6 HOLISTIC DEVOPS FRAMEWORK

At this point time has come to choose what style of development to have for each component. Requirement driven development has been the standard in ICS area. The one and only guiding star for many years has been CMMI certified, aiming on a maturity level 3 for commercial organisations. The question is whether that is the best for all anymore?

Looking at the figure 19 below should trigger us thinking of what the best development model for your product (Bosch, 2019, p. 163):
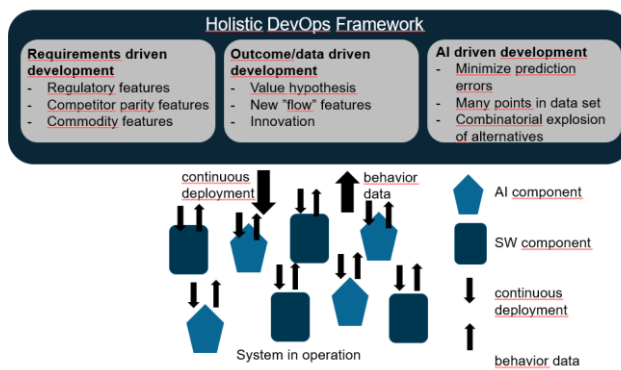

Figure 19. Holistic DevOps Framework **(Bosch, 2019, p. 163)**

Requirement driven development:

In traditional ICS and currently the only way of doing development in a safety domain. This way of doing development have many things loaned from *waterfall project management style* with a primary driver of getting "what you have ordered from a supplier is what you get". All requirements shall have defined how they should be validated. When changes come to a requirement, it needs to be updated and the whole change needs to be verified again. For further information into how to do requirements I suggest looking into NASA's guide for this (Hooks, 2000).

Outcome driven development

When desiring for an improvement in an area e.g. less false positives warning messages - assemble a cross functional team and give a goal to reduce with e.g. 10% in each increment until there is not more financial benefit aka it's good enough! The team will use standard tools like Ishikawa diagrams as normally the 80/20 rule (compare the Pareto principle) of the goals can be applied.

AI driven development

By collecting data from running installations and applying ML/DL in this data let the algorithms perform better according to one by us defined optimal outcome function!

What type of development is best for the product/software component for you one needs to be decided based on the constraints and system level requirements? If one has safety demands on some part that needs to be requirement driven - but that creates too much costs, it is suggested to limit down to areas where it's really needed. If one has a very well - tuned CMMI mature organization one can perhaps allow for its use more widely (own reflection Andreas Bäck 2019).

2.7 SUMMARY

Based on the theory chapter we see the multidisciplinary and its problems when trying to find a joint agreement on **what, when, where and how to go for a more digitalized organization**. Some paths are already tested and safe but there are new areas and buzz words with or without contents that make development actions more difficult.

I see that the building stones in theory *can be combined* to even better tools and this is what encourages me to carry out the study. It is my hope to improve things based on a more solid foundation.

Chapter 3 is the method chapter.

3 METHOD

Methods that will be used in this thesis consist of a **combination of many methods,** sometimes called triangulation see (Patel & Davidsson, 2010). I am an actor in the company, and I have studied literature as a foundation. I have also been part in workshops and read audit reports, top papers and gap studies done by external companies for Wärtsilä regarding safety, cyber security, lean, agile and autonomous shipping. During my thesis Wärtsilä has joined Software Center (Software Center, 2019) and that have given me access to the latest papers and research in this area when it is up to company´s internal research and publications.

During the process of this thesis I have also taken courses in most of the competence areas of the Micro drivers. Additionally I have taken several certificates in the areas concerned.

The dialogue with colleagues has been active during the whole process and we can call this a form of unsystematic and unstructured interviewing. These interviews and discussions have given me fruitful information both about actual literature, training courses, gurus in the area and a more critical skill to judge both details and hopefully to understand the bigger picture (compare Denscombe 2014).

In section 3.1 I argue about the validity of my study.

3.1 VALIDITY

The data needed for making decisions in this domain come from recommendations and demands written down in standards that either Wärtsilä have chosen to follow or that we must follow if we want to be in that market.

I mean that what I have presented in the theory chapter and the combination of methods in my study build up both the intern and extern validity. It is all about an attempt to develop one business area in a multinational company and I have chosen, as I see it, *the most fitted models*, referred to them and tried to understand them in, for me, an optimal way.

How data driven development ideally is introduced, described, explained and put in use is based on generally approved methods like DevOps, SAFe. The latest ideas come from papers and workshops that is an outcome from Wärtsilä having joined Software Center where companies and universities work and collaborate together to accelerate software engineering (Software Center, 2019).

3.2 CRITICS

Study is carried out in a volatile field. There are forces in the market which influence the customers decision making processes and there is of course tough competition and tough competitors plus a lot of ongoing research. Let us call this the MARKET!

There are several disruptive technologies and patterns ongoing in the world today that makes it very hard to predict the future. The truth yesterday is perhaps not the truth today or tomorrow. The oil price, the politics and the financial market can change over a night!

As an actor and at the same a part time researcher, it's important to continue gathering new info on both new technologies and methods, while keeping a lookout for changes and trends in the global market in order to make the right decision on what to develop to end up with the right business decisions (compare Denscombe 2014).

I have done my utmost to navigate between a full relativistic world order and 100 % rationality and I mean what I am presenting is a balanced insight for a company to use, to correct and to cultivate daily, monthly and yearly.

4 RESULTS

In this chapter I present the results of my study. The results consist of pictures, models, questionnaires, tools and advices that can be tested in the company in one way or the other. The results are based on theory building combined with observations, experiences learned, mistakes and lessons learned but also applying and refreshing of old ideas.

My results from my study are presented section wise from section 4.1. in this chapter, both as *descriptions, explanations and some normative advices*.

4.1 ECONOMICS AND REALITIES

Before going further its crucial to know what the resources we spend today are, and for how long a lifetime to go. I mean it's not worth the effort to change something on a product that soon will be obsolete. Looking at figure 20 I try to give an understanding of the commodity costs that are spent today on just maintaining a product.
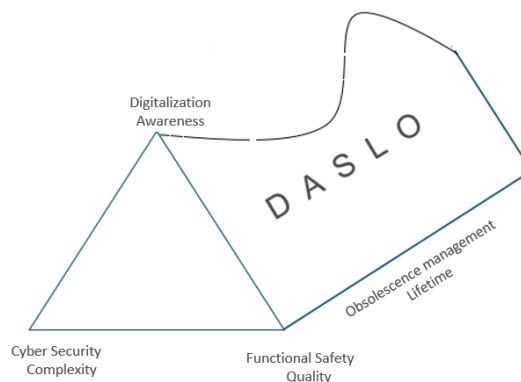

Figure 20. Micro driver's commodity costs model (Andreas Bäck 2019)

This model in figure 20 works as a model to visualize the cost of maintaining a product if the assumption is that all efforts are requirement driven and maturity models used are 1…5. on all the micro drivers into more generic ones enabling one to add more micro drivers in the future like "Open Source Software" (OSS) and "International Electrotechnical Commission for Explosive Atmospheres" (IECEx) (advice by Andreas Bäck 2019). I recommend anyone to familiarize DASLO and its contents!

Complexity is such things like micro drivers like c*yber Security* that will drive make the commodity baseline bigger. In this thesis we have talked about IEC 62443 and NERC-CIP. But OSS could be added into this later also.

Quality is all these things that make us have a higher quality as defined by a certain process. Here we start to see a shift in definition of quality of you compare CMMI defection and the more Agile if you look in chapter 2.2.

Lifetime is simply amount of years you plan to support this product or service. If one does not have planned the end days of the product it can be very expensive as the model is trying also to show.

Awareness is the micro driver *Digitalization* or any other means to needed and making your customer satisfied faster. The more you know about your customers business you can sell things to him that he really wants in the end when you know this very well a new business can be created out of the knowledge and data you have gathered.

The main idea with the DASLO model is to make it visible that when choosing higher maturity levels and lifetime the commodity costs can grow very big. Here it's crucial to select right levels to what the business is ready to pay for this.

4.2 PORTFOLIO PLANNING

Looking at the products in the portfolio it's crucial to know where the products and/or the services are in the lifecycle and what the possibilities are for data driven decision making in the real context when striving for the next level of B2D strategy. I have developed a questionnaire or "**card**" that can be used to find out where a product/service exists in both the portfolio and in the system alias IOT - Stack. The result is presented in table 3 as follows:

Table 3. Question card for identifying systems in an IOT-Stack - Andreas Bäck 2019

Question	Answer	Why
Product name		Unique name, see CPE in (NIST, 2019)
Product owner		Who is the owner of this and can answer the questions below as a GDPR owner?
Sellable product Yes/NO		Is this an inhouse system that is an enabler for a sellable product?
Stairway to Heaven Speed, Data, Ecosystem, AI		Understand digitalization maturity level. (Bosch, 2017, pp. 8-10)
Holistic DevOps Framework [Requirement, Data, AI]		How "Requirements" are managed. (Bosch, 2019, p. 163)
Analytics Cycle Time (Bergh, et al., 2019)		The time it takes from product is released and data from is taken into use in next release.
Quality Maturity Level No = 0. CMMI=1…5 (CMMI Institute, 2019)		Majority of products are developed under some quality system. ISO9001 ~ 1,5 (Paulk, 1994)
Safety device No = 0, SIL =1…5 (IEC 61508-*)		Can data be collected in the cloud? Do we need to have network segmentation? What level of quality standard to follow?
Firewall Segment 0= Safety Device (DNV)		Understanding of where in the IoT stack this device plays.
Cyber Security Level 1...4 No = 0 Yes = 1…4 SIL (IEC 62443)		Certified against IEC 62443?
Patch time in days (NERC, 2017)		Time after a CVE is official, and it's patched at customer. NERC - CIP 35 days
Autonomy levels AL0 … AL6 (Register, 2017)		What autonomy level in unmanned ship?
Biggest Pain to reach next AL level?		What is in the critical path for you to reach next Autonomy level?
Open Source Software maturity https://www.openchainproject.org or your survey answers from https://github.com/todogroup/survey		Are OSS components managed so one knows what CVE's one are affected by? Risk of being sued and bad publicity.

4.2.1 IOT - STACK ALIAS SYSTEM VIEW

When looking at the questions from table 3*, my suggestion for a brain storming session or workshop is to print out* table 3 *and place it on a whiteboard with e.g. background picture of a ship if that is what one tries to model* like in figure 21. From each card one should also draw lines on how the cards depends on each other connectivity wise, compare vectoring or sociometry!
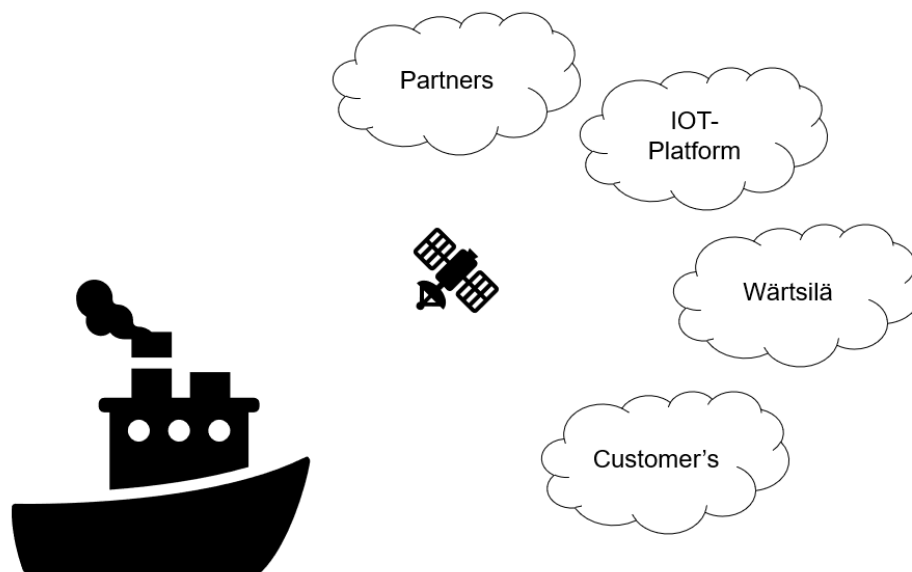

Figure 21. Canvas for planning an IOT- Stack with Table 4 questions - Andreas Bäck 2019

4.2.2 PORTFOLIO VIEW

Referring to the table 3 "**Card**" my advice is now to go to a portfolio view. In this view one is more interested in getting a picture of in what state of the lifecycle the product/service is and if we should invest in getting data out of it. Placing all products in Ansoffs P/M matrix (Wadström, et al., 2017, p. 84) like in figure 22 will give us insight in to the portfolio! The matrix is presented and described in figure 22 as follows;
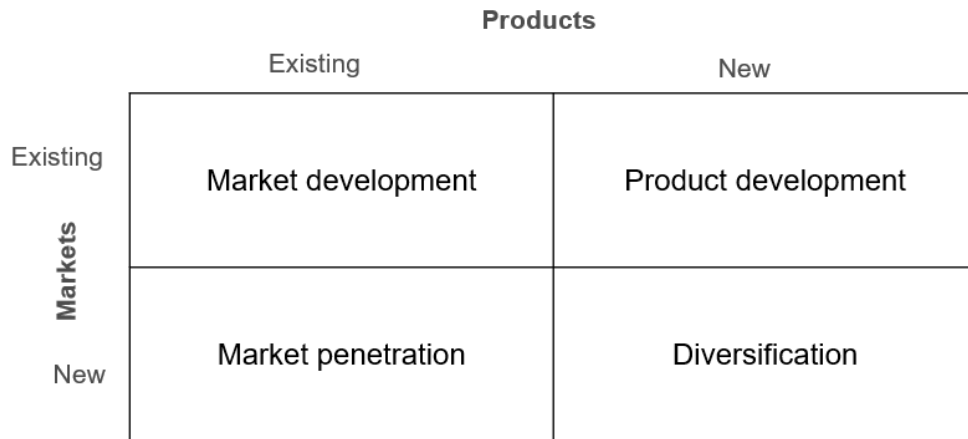
Products

	Existing	New
Existing	Market development	Product development
New	Market penetration	Diversification

(Markets — vertical axis: Existing, New)

Figure 22. Ansoffs P/M matrix. **(Wadström, et al., 2017, p. 84)**

When deciding on what product to really invest in both in a business sense and a data driven sense, I recommend the matrix developed by Boston Consulting Group as a helpful tool to be put into use! (Wadström, et al., 2017, p. 119). Crucial questions to get answers to are as follows:

- Do we have overlapping products/solutions in the portfolio?
- How is the lifecycle stages alias DASLO (see table 2 in this study/Andreas Bäck 2019) overlapping with each other?
- Do we have enough incentives to spend money on data driven attempt for this product/service?

When the strive is to deepen the analysis for a more critical but also balanced decision making my advice is to also use an BCG matrix approach. In business theory and in mathematics we sometimes call this type of actions for BOOT STRAPPING! The BSG matrix is described as follows:
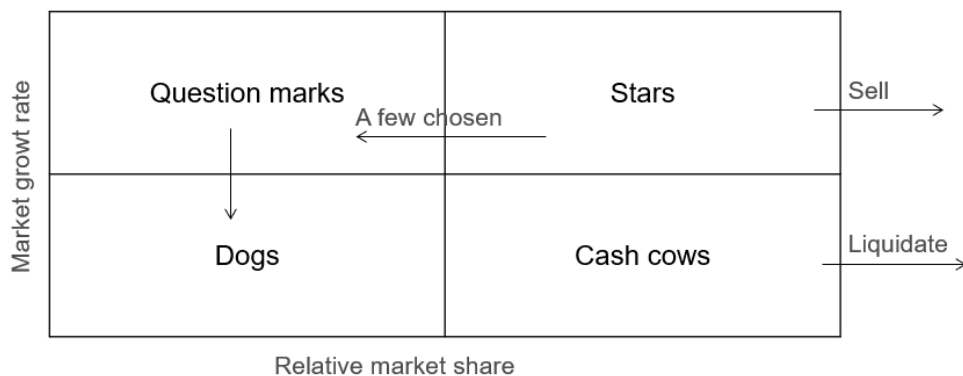
Market growt rate (vertical axis) / Relative market share (horizontal axis)

Question marks	Stars	Sell
Dogs	Cash cows	Liquidate

A few chosen

Figure 23. Boston Consulting Group matrix (BCG-Matrix) **(Wadström, et al., 2017, p. 119)**

4.2.3 VALUE STREAM MAPPING

The next advice in my process guidance is to consider the value stream mapping findings.

A value stream mapping done in parallel with the questions in table 3. will help the work when you have decided that you will start the digital change culture – somewhat of a first GREEN LIGHT is hopefully there! As there will be major changes in both how one works with technology and changing into cross functional teams in order to have higher business agility and making customers more satisfied *value stream mapping* serves as an identifier of dependences and reveals waste, unnecessary!

The important question is "how can we reduce the cycle time of our product or system?" also mentioned as speed in STH, Stairway to heaven (Bosch, 2017, pp. 8-10). The answer to this important question will lead you/us right.

4.2.4 PRODUCT REGISTER

When the first iterations of the "Product cards", see table 3, are done it is recommended to create a register of all this as storing the information is something that is needed for a long time forward. Why so? Yes – "we must know" …

- Who is the owner when there is an incident or compliancy question?
- On what level and how is a product managed?
- What data is possible to share and combine/normalize and who owns it?

4.3 CATEGORIZING DATA AS A CDO

As we have seen above, we need to create good business packages that delivers value all the time so we can continue with more Minimum Viable Products (MVP). By combining several of the tools the following questions in table 4 are there to help us find the right mix of data, so the business has both a long - term and a short - term profit basis. The need for categorization of data is due to the need for useful data and the tough business-related requirements and facts to sort out things - variables that do not explain anything or explaining almost nothing. The tool for categorizing is as follows in table 4:

Table 4. Decision tool for selecting right data projects to work - Andreas Bäck 2019

Question	Why?
What is the unique name of data? See CPE in for the intention (NIST, 2019)	Unique name to avoid duplicates.
User story?	Understand what is requested. As a <role> I can <capability>, so that <receive benefit>
[Structured| Unstructured| Semi-structured| Meta| None] data today? (Erl, et al., 2015, p. 17)	To understand complexity with this data.
[Requirement| Outcome| AI] intended driven development?	In what way will the lessons learned from the data be inserted back into the development?
[Offensive| Defensive] Data Strategy? (Bergh, et al., 2019, p. 59)	How will we divide our Data Strategy portfolio?
[Commodity| Differentiation| Innovation] development? (Bosch, 2018, p. 96)	One would like not to spend all effort on Commodity.
Time before transition to previous level in 3PLM (Bosch, 2018, p. 96)	Understanding time critically before this data becomes a Commodity
How much CAPEX or resources are needed before platform ready to deliver first data?	How much money, time and resources this development project will require for an MVP?
How much OPEX needed to maintain data after we have implemented this.	Understand if we need to have personal updating requirements, fees for cloud computing in an AI development.
How many organizational transactions or silos need to be involved?	Is an organizational change needed according to BAPO and creation of cross functional teams

4.3.1 APPLYING QUESTIONS

The following activity in my process guidance is about applying questions in order to select out the combination of data that fits our strives in the "best way" or/and the most fitted model paradigm. At this point one should have a list of all the data initiatives and now it's up to decide what will be done and in what order with the budget you have. My suggestion is to use the tools

described in this thesis like "Ansoffs P/M matrix" and define guardrails. Guardrails could preferably be Commodity (70 %), Differentiation (20 %), Innovation (10 %) (Bosch, 2019, p. 76) according to 70/20/10 model based on 3 Horizons. These distributed numbers should be aligned with where You are today and where You would like to go. Inside the guardrails my suggestion is to use tools like "Weighted Shortest Job First" (WSJF) (© Scaled Agile, Inc., 2018).

One could also use the "Data strategy" with [Offensive| Defensive] found in table 4 for selecting what is the right balance of data. Wärtsilä is in my opinion somewhere between Hospitals and Banks if one looks on figure 24 and this would give a balance around 60/40.
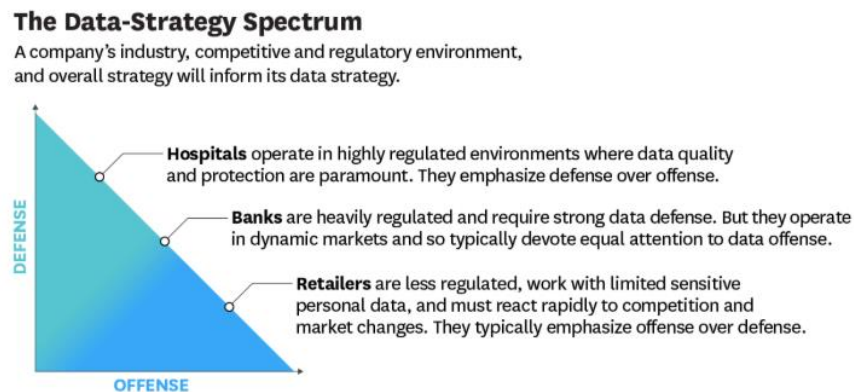
Figure 24. What is your Data Strategy? (**Davenport, 2017**)

4.3.2 CULTURE CHANGE TOWARDS DIGITAL TRANSFORMATION

When the initial steps in the guidance process are passed there is perhaps a need and added insight and maturity available to *head for a culture change*. Digital transformation is unfortunately not so well served by the pure Project Oriented Management (Kersten, 2018, p. 54) approach when referring to (Kersten, 2018).

My advice when wrestling with a B2D transformation is both to combine (Kersten, 2018) and (Bosch, 2019) insights and findings. Bosch´s findings are more about business and Kersten is more oriented towards and concerned about architecture, processes and activities in a business context, yes. (Bosch, 2019, p. 5) In Figure 25 we see how to prepare and what to start from when striving for a B2D journey as follows:
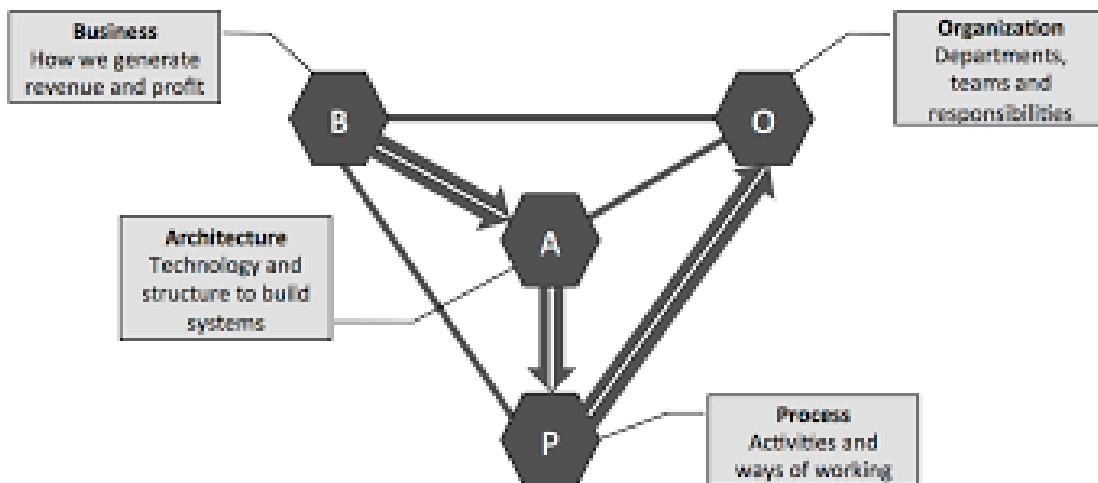
Figure 25. The BAPO model (**Bosch, 2019, p. 190**)

The strength with this model is that one builds the strategy, the processes and the activities based on business needs meaning that CNM (Customer Needs Management, see www.gartner.com 2019) is not forgotten – on the contrary – highly prioritized which both research and practice loudly call for!

4.3.3 CROSS FUNCTIONAL TEAMS

Additionally, to descriptions, explanations and advices presented so far, I finally present a tool to assist in transforming to cross functional teams. When doing a change from functional teams to cross functional teams it can be very challenging to get there in one go. So below, I have developed a model called "Tower of Speed". figure 26 is a model on how one can go from "functional" silos into cross functional teams in stepwise enabling teams to deliver business differentiators in one "floor change" alias commence addition in the tower. The tool is as follows:
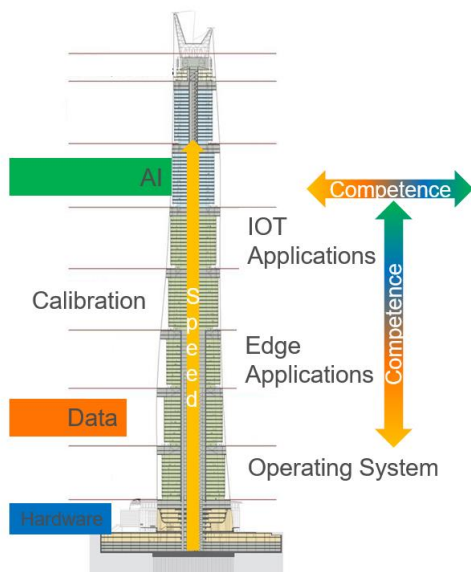
Figure 26. Tower of Speed tool - Andreas Bäck 2019, *own version* based on collected information.

Ideas that makes up this tower comes from:

- The SAFe House of Lean. (© Scaled Agile, Inc, 2018)
- Stairway to Heaven, (Bosch, 2017, pp. 8 - 10) see the color coding in Figure 12.
- The Three Layer product model. (Bosch, 2018, p. 96)
- T - Shaped people. (Wikipedia, 2019)
- Lean flow.
- Persons being in the zone alias flow (Csikszentmihalyi , 1997, p. 31)
- One can't get from ground floor to the top of Shanghai tower with one elevator!

To in one big step go from an expert in a narrow are and learn a wide are of knowledge in one go is very challenging and perhaps not all have to know the full "Tower" knowledge span either. Ideas comes from that people feel best about them self when little bit outside of comfort zone but not too much. This model also includes elements so that all people in one step could be an apart of delivering solutions that are differentiators and are healthy for the company future. Important in this model is also to foster teamwork, empowered teams and enabling people to grow into architects and product owners. The color coding is an addition to the Stairway To Heaven that we in Wärtsilä have applied to clearer indicate what area we talk about.

Declaration behind the idea is as follows:

Benefits of having cross functional teams to foster T - Shaped people are quite many and it's first when you start to work in this way many of these things becomes obvious and the waste is found. Benefits according to my summarizing based on literature are as follows:

1. When organizations are merged waiting time between them will be reduced.
2. New skills are learned from each other and problem solving get new tools.
3. New eyes can question why we are doing this. The 5 Whys Process is a good tool from six sigma and lean.
4. All will work towards same target and different agendas will be reduced.
5. Tacit information will be documented.
6. Organization will broader competence and can react better to changes on demands.
7. More resilient to bus factor (Fitzpatrick & Collins-Sussman, 2012, p. 7)

(Benefits with T-Shape and cross functional teams according Andreas Bäck 2019 – Master´s thesis)

4.4 MICRO DRIVER BENEFITS WITH DATA DRIVEN DEVELOPMENT

When linking to the introduction in chapter I here explain and summarize micro driver benefits with data driven development. The drivers presented are cornerstones for any company when planning for their own digital transformation.

Safety micro driver in this thesis that have least benefit from a data driven approach to development. It's even so that it's difficult today to have too much Agile workings in this environment as all standards expect and validates against requirement driven development. Most of the benefit of data driven here is the collection of evidence that processes have been fulfilled but that does not require a closed looped of data, it can even be challenging to get data as Safety devices have regulation on network segmentation and this in practice prohibits data collection in the production environments.

Lifetime is the micro driver in this thesis that defines the lifetime budget from a data driven way of development.

- There are profits to be made in the end of a product if planned right as customers will do last time buys instead of cost due to neglected planning.
- Possibility to do the right decisions and plan when product should be not developed and more and resourced directed to where there can be made profit.
- Alignment of products in a system to enable offerings few other competitors can do.

Cyber security is impossible to do in the long run without releases made on demand - how else could one fulfill demands like NERC CIP demanding a mitigation time of 35 days? Now this domain majority is mostly driven by regulations and that is more driven by processes than by speed.

- Up to date cybersecurity and improved quality and availability due to fast deployment of security patches and bug fixes. (AG, 2019, p. 4)

Digitalization is a combination of all the tools (Agile, DevOps, DataOps, Lean, Six Sigma and Continuous Delivery) and technologies combined in order to produce new digital offerings. It's impossible without:

- Fast delivery of customer value and integration of customer feedback through delivery in much shorter cycles than today. (AG, 2019, p. 4)
- Continuous, data driven increase of value throughout complete product life time due to enhancements based on operational data and A/B field experiments. (AG, 2019, p. 4)**.**
- Reduced deployment and operational risks through highly automated delivery of small changes. Vision: deployment is a "non event". (AG, 2019, p. 4)

4.5 SOME PRACTICAL CONTRIBUTIONS TO THE COMPANY

Here as a summary I suggest four advices or ways for improving the started digitalization process in the company as follows:

1. System level portfolio management (AG, 2019) earns more priority and efforts!
2. Network value stream mapping and creating cross-functional ownership, architecture and processes on that into focus!
3. Involve Top Management and buy in their support for and onto the B2D journey! We have all to be committed to the infinite game!
4. Get experienced people on board that have been on an agile transformation before to act as senior coaches and give encouraging advices "in the darkest hours, glimpse of hope I mean!"

If a company does not actively start the transformation process there is a risk the process will never really start, it's also challenging to find the right time to do this so just start. Let us quote Martin Luther King. 1611 that said: "The time is always right to do the right thing."!

To assist to take the right decision for digitalization have for me at this point been the question. *Will this make us faster in the future?* This simple question is my current compilation that what digitalization means for me. I will below give light into why this question works as a good guiding star when making decisions. Below follow some sentences that explain my thoughts to this question.

- More often we have to do more with less.
- With a smaller R & D budget it is often hard or almost not easy to "gear up"
- It's the automation people that usually must be the forerunners in industrial control systems, so we must show with example even if it can be hard.
- Technical depth has to be fixed alongside normal development
- Fast feedback make innovation happen faster and people motivated.

(Toughs that made up question "Will this make us faster in the future?" Andreas Bäck 2019)

5 CONCLUSION

In this chapter I am summarizing the results from the study. In section 5.1 I ask whether I reached the purpose. Section 5.2. is about contributions and in section 5.3 there are proposals for further research.

5.1 HOW DID I REACH MY PURPOSE?

The primary purpose was to give advice about what improved data quality can do in an industrial control ecosystem. I have presented descriptions, explanations and advices on how to think on the more regulatory domains and the constraints.

The secondary purpose was to come up with descriptions, explanations and advices how to work with data. Here I developed several tools that help to decide what products/services and systems should get selected to the digitalization strategy.

Based on my study and results obtained I consider that the purposes of the study have been successfully reached.

5.2 MY CONTRIBUTION

My contributions are two folded. *Firstly,* I have gathered a lot of literature and theories and sorted out the golden nuggets. This study has enabled me to have the opportunity to go through much more information than ever before and that is rarely possible to do in the normal work life. *I have learnt to learn*! It is a hope that we in Wärtsilä to some extent also will apply theories and some of my guidance in our industry field.

Secondly several tools were created and listed both in tables and figures and equipped with questions for deepening practical actions. All the tools have their ideas that connect back to other known tools and theories. I recommend that the company starts to use some of the tools now, together we can adjust them and use the as the most fitted ones in our case for our needs and wants.

5.3 PROPOSALS FOR FURTHER RESEARCH

Based on this study and reflections I present three proposals for further research as follows:

1. Current budgeting models for automation where R&D gets a percentage of EBIT isn't optimal in the future, to investigate on how models for this this could be done with help pf Product Lifecycle Management (PLM). Current trends indicate that:

2. Investments in hardware and electronics will be lower, there is predictions that hardware will only be the cost of filament with additive manufacturing.

3. Digitalization will grow the revenue on products through digital offerings and here software, data and AI is the key, and this will lead to that automation will take a bigger part for the R&D budget, if not also from other budget streams?

4. New version of Obsolescence management 62402:2019 become available while being on the last workings on this thesis. I have used 62402:2007 in this thesis. New one has more lifecycle phases and written as a requirement instead of guideline so would be useful how this can now better be incorporated in the processes. And software was also mentioned in this new standard that was totally lacking in the past.

Weighted Shortest Job First (WSJF) model could be further developed to consider things that makes us faster, prioritize differentiators and innovators and work that will pay back for a long time e.g. on things that will evolve into new digital offerings.

Setting up an AI driven platform that allows a rapid data cycle in industrial systems. One can't just copy the concept from domains that are today in the forefront, Energy and Marine sector have many constraints like regulations, embedded systems and a very long product lifecycle.

Investigate on how the definition of quality could be improved in the areas of regulations as it still very much depends of standards that assume quality is measured by processes and paperwork produced and not really that customer is satisfied.

6 REFERENCES

© Scaled Agile, Inc., 2018. *Weighted Shortest Job First.* [Online]
Available at: https://www.scaledagileframework.com/wsjf/
[Accessed 29 09 2019].

© Scaled Agile, Inc, 2018. *lean agile mindset.* [Online]
Available at: https://www.scaledagileframework.com/lean-agile-mindset/
[Accessed 3 12 2019].

ABB Drives services, 2016. *EN_Services_catalog_RevB_3AUA0000182592.pdf.* [Online]
Available at: https://search-
ext.abb.com/library/Download.aspx?DocumentID=3AUA0000182592&LanguageCode=en&
DocumentPartId=1&Action=Launch
[Accessed 6 01 2018].

AG, S., 2019. *Continuous Delivery and DevOps,* s.l.: Siemens IoT Services.

Andersson, C., 2015. *Creating a Data-Driven Organisation.* 1st ed. Sebastopol: O´Reilly.

Anon., 2017. *Quality (business) - Wikipedia.* [Online]
Available at: https://en.wikipedia.org/wiki/Quality_(business)
[Accessed 04 08 2017].

Arbnor, I. & Bjerke, B., 1994. *Företagsekonomisk metodlära.* Lund: Studentlitteratur AB.

AXELOS, 2019. *ITIL.* [Online]
Available at: https://www.axelos.com/best-practice-solutions/itil
[Accessed 16 11 2019].

Beck, K. et al., 2001. *Manifesto for Agile Software Development.* [Online]
Available at: https://agilemanifesto.org/
[Accessed 7 12 2019].

Bergh, C., Benghiat, G. & Stroud, E., 2019. *The DataOps Cookbook.* 1st ed. Cambridge:
United States of Amerca.

Bosch, J., 2017. *Speed, Data, and Ecosystems: Excelling in a Software-Driven World.* 1st ed. Boca Raton: CRC Press .

Bosch, J., 2018. *Impactful Software.* 1st ed. United States of America: Kindle Direct Publishing.

Bosch, J., 2019. *Digital Transformation A Holistic Perspective for Business Leaders.* 1st ed. s.l.:s.n.

Bosch, J., 2019. *WASP FACULTY AND INDUSTRY DAYS.* [Online]
Available at: https://www.youtube.com/watch?v=MTt30zyd38w&t=23s
[Accessed 3 11 2019].

Bruzelius, L. H. & Skärvad, P., 2017. *Integrerad organisationslära.* 11 ed. s.l.:Studentlitteratur AB.

Brynjolfsson, E., McAfee, A. & Brynjolfsson, E., 2014. *The Second Machine Age.* 1st ed. New York : W. W. Norton Co.

CMMI Institute, 2019. *CMMI® Services.* [Online]
Available at: https://cmmiinstitute.com/cmmi/svc
[Accessed 6 12 2019].

Csikszentmihalyi , M., 1997. *Finding flow.* 1st ed. NewYork: Basic Books.

DataKitchen, 2019. *The DataOps Manifesto.* [Online]
Available at: https://www.dataopsmanifesto.org/
[Accessed 7 12 2019].

Davenport, T. H., 2017. *whats your data strategy.* [Online]
Available at: https://hbr.org/2017/05/whats-your-data-strategy
[Accessed 6 12 2019].

DonkeyCar, 2019. *donkeycar.* [Online]
Available at: https://www.donkeycar.com/
[Accessed 16 11 2019].

Erl, T., Khattak, W. & Buhler, P., 2015. *Big Data Fundamentals.* 1st ed. Crawforsdville: ServiceTech PRESS.

European Union, 2017. *Home Page of EU GDPR.* [Online]
Available at: http://www.eugdpr.org/
[Accessed 31 07 2017].

Fader, P. S., 2012. *Customer centricity : focus on the right customers for strategic advantage.* 2nd ed. Philadelphia : Wharton Digital Press.

Fingrid, 2017. *Welcome - Fingridin avoin data.* [Online]
Available at: https://data.fingrid.fi/en/
[Accessed 03 08 2017].

Fitzpatrick, B. & Collins-Sussman, B., 2012. *Team Geek.* 1st ed. Sebastopol: O'Reilly Media.

Gartner, 2017. *From the Gartner IT Glossary.* [Online]
Available at: https://research.gartner.com/definition-whatis-digitalization
[Accessed 25 07 2017].

goleansixsigma, 2019. *8 Wastes Check Sheet.* [Online]
Available at: https://goleansixsigma.com/the-8-wastes-checksheet/
[Accessed 07 12 2019].

Günther, P. A., 2017. *The Blue Economy version 2.0.* 1st ed. New Delhi: Academic Foundation 2015.

Hooks, I. F., 2000. *Guide for managing and writing requirements.* Boerne: Requirement Experts.

IEC 61508, 2010. *IEC 61508 Explained.* [Online]
Available at: http://www.iec.ch/functionalsafety/explained/
[Accessed 25 07 2017].

IEC, 2010. *IEC 61508-3:2010,* s.l.: International Electrotechnical Commission.

Institute, M. G., 2011. *1.* [Online]
Available at: https://www.mckinsey.com/featured-insights/urbanization/urban-world-

mapping-the-economic-power-of-cities

[Accessed 14 9 2019].

ISO/IEC 9001, 2015. *ISO/IEC 9001:2015.* s.l.:IEC.

Kahneman, D., 2013. *Thinking, Fast and Slow.* 1st ed. New York: Farrar, Straus and Giroux.

Kersten, K., 2018. *Project to Product.* 1st ed. Portland: IT revolution press.

Knapp, E. D. & Langill, J. T., 2015. *Industrial Network Security.* 2nd ed. Waltham: Syngress.

Lwakatare, L. E. et al., 2019. *A Taxonomy of Software Engineering Challenges for Machine Learning Systems: An Empirical Investigation.* Gothenburg, SpringerLink.

Michaelides, M. P., Herodotou, H., Lind, M. & Watson, R. T., 2019. *Port-2-Port Communication Enhancing Short Sea Shipping Performance,* s.l.: MDPI.

NERC, 2017. *CIP Standards.* [Online]
Available at: https://www.nerc.com/pa/Stand/Pages/CIPStandards.aspx
[Accessed 6 12 2019].

NIST, 2019. *Security Content Automation Protocol.* [Online]
Available at: https://csrc.nist.gov/projects/security-content-automation-protocol/scap-community
[Accessed 6 12 2019].

Patel, R. & Davidsson, B., 2010. *FORSKNINGSMETODIKENS GRUNDER.* 4:th ed. Lund: Studentlitteratur AB.

Paulk, M. C., 1994. *A Comparison of ISO 9001 and the Capability Maturity Model for Software.* [Online]
Available at: http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=12187
[Accessed 1 5 2019].

Project Management Institute, 2019. *Project Management Professional (PMP)®.* [Online]
Available at: https://www.pmi.org/certifications/types/project-management-pmp
[Accessed 26 10 2019].

Redman, T. C., 2008. *Data Driven: Profiting from Your Most Important Business Asset.* 1st ed. Boston: Harvard Business Review Press.

Register, L., 2017. *Design Code for Unmanned Marine Systems..* [Online]
Available at:
https://www.cdinfo.lr.org/information/documents/ShipRight/Design%20and%20Construction/Additional%20Design%20Procedures/Design%20Code%20for%20Unmanned%20Marine%20Systems/Design%20Code%20for%20Unmanned%20Marine%20Systems,%20February%202017.pdf

Rosenberg, J., 2014. *Twitter.* [Online]
Available at: https://twitter.com/jjrosenberg/status/502466278221561857
[Accessed 04 08 2017].

Runeson, P. & Höst, M., 2008. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering,* Volume 14, pp. 131-164.

Scaled Agile, Inc., 2018. *Identify Value Streams and ARTs.* [Online]
Available at: https://www.scaledagileframework.com/identify-value-streams-and-arts/
[Accessed 28 09 2019].

Skärvad, P.-H. & Olsson, J., 2017. *Företagsekonomi 100 Fakta.* 18th ed. Solna: Liber.

Smith, D. J. & Simpson, K. G., 2011. *Safety Critical Systems Handbook.* 3:rd ed. Burlington: Butterworth-Heinman.

Software Center, 2019. *Software Center.* [Online]
Available at: https://www.software-center.se/
[Accessed 15 10 2019].

Software Center, 2019. *SOFTWARE CENTER.* [Online]
Available at: https://www.software-center.se/
[Accessed 26 10 2019].

Software Engineering Institue, 2010. *CMMI for Services, Version 1.3,* Hanscom: Carnige Mellon.

Ståhl, D. & Mårtensson, T., 2018. *Continues Practices.* 1st ed. Milton Keynes UK: Lightning Source UK Ltd..

Statista, 2017. *Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions).* [Online]
Available at: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
[Accessed 25 07 2017].

Wadström, P., Schriber, S., Teigland, R. & Kaulio, M., 2017. *Strategi.* 1st ed. Stockholm: Liber.

Wärtsilä , 2019. *Smart Marine.* [Online]
Available at: https://www.wartsila.com/marine/smartmarine
[Accessed 3 12 2019].

Wartsila, 2017. *next-generation-unic-automation-system-enables-wartsila-31-performance.* [Online]
Available at: https://www.wartsila.com/twentyfour7/in-detail/next-generation-unic-automation-system-enables-wartsila-31-performance
[Accessed 25 07 2017].

Wärtsilä, 2017. *Wärtsilä 31 is recognised by Guinness World Records.* [Online]
Available at: http://cdn.wartsila.com/docs/default-source/product-files/engines/ms-engine/brochure-o-e-w31.pdf
[Accessed 25 07 2017].

Wärtsilä, 2017. *Wärtsilä Brand Hub - Home.* [Online]
Available at: https://www.wartsila.com/brandhub
[Accessed 03 08 2017].

Wärtsilä, 2019. *Smart Marine.* [Online]
Available at: https://www.wartsila.com/marine/smartmarine
[Accessed 7 12 2019].

Wikipedia, 2019. *T-shaped skills.* [Online]
Available at: https://en.wikipedia.org/wiki/T-shaped_skills
[Accessed 03 12 2019].

Wiseman, R., 2004. *The Luck Factor.* 10th ed. London: Arrow.

7 TABLE OF TABLES

8 TABLE OF FIGURES