Accelerate Business Value in Manufacturing with Advanced Analytics

Axel Lundholm

Lund University
Department of Automatic Control

MSc Thesis
TFRT-6089
ISSN 0280-5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by Tryckeriet i E-huset
Lund 2019

Abstract

Recent developments in smart manufacturing enable convergence between the digital and physical worlds of modern manufacturing facilities. This evolution is, however, far from trivial and thorough research and investigation needs to be conducted regarding dynamic connectivity of assets and implementation of data driven analytics, which provides deeper insight into the operational processes.

Scania in Södertälje is the object for the case study, with the aim of presenting recommendations for future research projects within smart manufacturing. Also, PTC in Boston, Massachusetts, has contributed with expertise and knowledge in the matter. Addressing the problems regarding what future actions to pursue and what methodologies to invest research in, this thesis base its analysis and discussion on recent research papers and documentation from international standardization organizations.

The analysis identified and categorized the present problems into *company wide development architecture*, *information modeling*, *communication structures*, *computational modules* and *collaboration with other companies and organizations*.

Ultimately, four different project recommendations are presented. The first suggestion includes development of a framework for using Reference Architecture Model Industrie 4.0 (RAMI 4.0) in company wide development and research, as a way of categorizing systems and functions. Secondly, a suggestion for generic modeling of assets was presented. Assets could be anything from machining tools, to analytics software and even operational personell. Thirdly, the thesis recommends that Scania investigates dynamic communication structures, which breaks the traditional hierarchical view on information infrastructure across the company. Lastly, a project regarding non-intrusive, online computational modules was discussed. This suggestion was, however, not in particular detail, as the thesis concluded that the foundation for data driven methods is of the highest importance, rather than suggesting actual analytics algorithms.

Keywords: *Smart manufacturing*, *Industrie 4.0*, *Reference Architecture Model Industrie 4.0*, *Asset Administration Shell*, *Cyber-Physical Systems*, *Big Data*, *Advanced Analytics*

Acknowledgements

First and foremost, I want to thank my advisor Prof. Charlotta Johnsson at the Department of Automatic Control at the Faculty of Engineering, Lund University. Your professional ideas combined with great mentorship have proven vital for the completion of this thesis work.

I also want to send my regards towards Scania in general and my advidor David Chauca in particular for making this thesis work possible. Furthermore, I want to thank all colleagues I have met at Scania, who have provided me with insight and experience.

At PTC, I initially want to express my thanks towards Hans Nilsson, for your extended efforts in initiating this project and introducing me to PTC and the fantastic city of Boston. This especially includes Dr. Jordan J. Cox and the team at Global Academic Programs, who welcomed me with open arms into not only a new city, but a completely new country.

Furthermore, I want to thank my examiner, Prof. Anders ”Robotsson” Robertsson, for everything you have taught me and for the inspiration you have given me.

Following this, I want to extend my thanks to my father, Dr. Thomas Lundholm. Without your vast knowledge and deep experience, this thesis work may never have been accomplished.

I also want to express my gratitude towards my mother, Annika Brommesson, for your never ending support and important thoughts on how to learn from difficult situations.

Last, but certainly not least, I want to direct my sincerest thanks to the great love of my life, Erica Juhlin. Your unmatched energy and inexhaustible faith in me has contributed to making this year in Boston and Lund one that I will never forget. This may be the end of our lives as students, but surely the beginning of an incredible new era.

Thank you all!

Contents

List of Figures

List of Acronyms

I4.0	Industrie 4.0
IoT	Internet of Things
CPS	Cyber-Physical System
CPPS	Cyber-Physical Production System
BD	Big Data
AA	Advanced Analytics
ML	Machine Learning
ISO	International Organization for Standardization
IEC	International Electrotechnical Commission
CP	Communication and Presentation
RAMI 4.0	Reference Architecture Model Industrie 4.0
AAS	Asset Administration Shell
OPC UA	Open Platform Communications Unified Architecture
AutomationML	Automation Markup Language
umati	universal machine tool interface
EBC	Equipment Behavior Catalogues
ERP	Enterprise Resource Planning
MES	Manufacturing Execution Systems
SCADA	Supervisory Control And Data Acquisition
PLC	Programmable Logic Controller

1

Introduction

This chapter will introduce the reader to the incentives of this thesis. It will also present some delimitations for the study and state the research questions. Lastly, a short outline of the thesis will be introduced.

1.1 Motivation

"Data is the new oil!", the UK Mathematician Clive Humby declared in 2006 [Palmer, 2006], *"It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value."*, implying that huge potential lies in the development of data driven approaches.

Industrial companies are in the process of implementing Industrie 4.0 strategies in their factories in order to realize better manufacturing efficiency and higher product quality. Extensive amounts of data are gathered from different equipment in modern manufacturing facilities and are output through multiple information channels and used in different systems and applications. With new disruptive technological advancements comes the need for a more systematical approach for gaining insight and knowledge from the manufacturing and enterprise data, in order to further automate the decision making process in industrial facilities.

The financial gain of Industrie 4.0 is hard to predict, but the general consensus is that the impact will be significant. In a report from Boston Consulting Group from 2015 it is suggested that Industrie 4.0 will contribute about 1 % per year to Germany's GDP over ten years. This includes creating up to 390,000 jobs and adding €250 billion to manufacturing investments in Germany alone [Rüßmann et al., 2015]

In order to further investigate the process of migrating into smarter manufacturing, a collaboration between Scania in Södertälje and LTH has been initiated. This collaboration will review the recent developments in the industry and suggest future lines of action for proceeding with smart manufacturing.

Scania's ambition is to start using acquired data in order to create business value for their manufacturing and logistics areas, as a part of implementing the fourth industrial revolution. Scania, amongst many other companies, is investigating how Industrie 4.0 and Advanced Analytics could provide insight, improve and accelerate production across the entire company. This thesis will base this research on Scania's internal goals for safety, quality, production efficiency and throughput.

It takes great effort for a factory to migrate over to a fully connected, Industrie 4.0 type operation. Rather than addressing all of the areas associated with Industrie 4.0, the ambition with this project is to try to address areas such as downtime prevention, product quality assurance, production volume increase from the perspective of Cyber-Physical Systems, Big Data and Advanced Analytics.

1.2 Scope

1.2.1 Aim and delimitation

The aim of this project is to investigate the future of the manufacturing industry in terms of connectivity, analytics and automation, with emphasis on developments to increase value for manufacturing businesses. Ultimately this thesis will serve as a guide or roadmap for how to prepare and adapt factories for the fourth industrial revolution, based on modern research in the field.

As this thesis takes a broad and strategical perspective, it will not discuss the theory behind particular algorithms, data structures, etc in depth. It will instead map out different approaches to how proceed with Industrie 4.0 and how to implement data driven Advanced Analytics into the business' operation and decision making.

Moreover, because of the thesis work's investigative nature and the relatively broad perspective on theoretical areas, all aspects presented in the theory chapter may not be of equal relevance in the analysis and discussion.

1.2.2 Problem formulation

This thesis specifically aims to answer the following questions:

- What actions should be performed to drive the integration of Advanced Analytics into a manufacturing workflow?
- What would be the ideal Advanced Analytics methodologies for usage in different manufacturing applications?

Based on the research made in this thesis, Scania IT could initiate future projects with more focus on actual data processing and analytics.

1.3 Outline

This section will briefly present the structure of the thesis.

Chapter 2 will introduce the approach used in the thesis work as well as the project process. Chapter 3 will present the necessary theoretical aspects for Industrie 4.0, divided in three categories; namely Cyber-Physical Systems, Big Data and Advanced Analytics. Chapter 4 will then introduce some background information regarding the case at Scania used in this study. Chapter 5 presents the main analysis of the thesis, where the theoretical aspects from the literature study are applied to the case. This chapter will try to answer the research questions that are presented in section 1.2.2 and present recommendations on future projects. Chapter 6 will hold the discussion of the findings made in previous chapters and evaluate how well the research questions were answered. Chapter 7 will, eventually, conclude the thesis along with the suggestions on future research topics made in the analysis.

2

Approach

This chapter will present the approach used in the thesis, as well as describe the project process – from the project initiation to the final analysis of results.

2.1 General methodology

As mentioned in section 1.2.2, the thesis will recommend future lines of action in the area of smart manufacturing. This will be done using an investigative approach. First, a general background covering the different ideas of smart manufacturing was methodically screened. From this, an extensive literature review was conducted; covering multiple aspects within the fields of Cyber-Physical Systems (CPSs), Big Data (BD) and Advanced Analytics (AA). Following this, the literature study was reassessed, to ultimately form a number of concrete recommendations for future projects at Scania. The process is illustrated in figure 2.1, and is explained in more detail in the following sections.

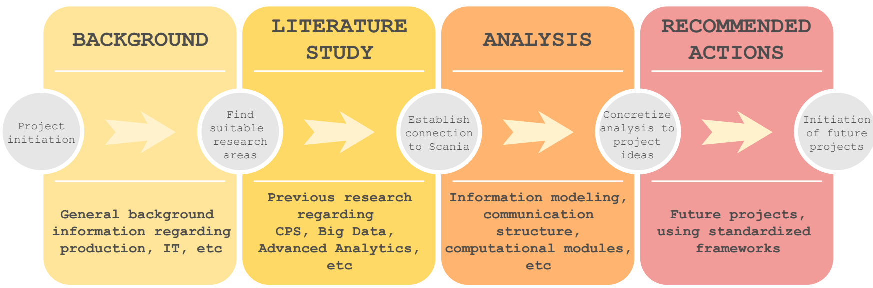

Figure 2.1 Approach for the thesis work

2.2 Background

In the initiation phase of the project, the areas of interest were Industrie 4.0 and Internet of Things. The first episode of the thesis work was conducted in collaboration with the U.S. company PTC, at their offices in Boston, MA. In order to gain relevant knowledge about Internet of Things (IoT) in general and the practice of IoT modeling in particular, a number of courses were taken at PTC, along with discussions with the supervisor at PTC and other experts on site.

Simultaneously, discussions were conducted with the Scania supervisor about how to apply the gained knowledge to Scania's manufacturing routines. As the project itself was not requested by a manufacturing division within Scania, but from the Smart Factory department in the IT section, part of the work became to find a suitable area to analyze from a theoretical perspective.

The initial idea was to use the massive amounts of data from Scania's top modern Body in White (BiW) factory in Oskarshamn, together the assistance from manufacturing experts on site solve to a problem present in the factory. This would be accomplished with the tools and expertise available at offices of PTC. Due to an aggregated workload within the Oskarshamn factory, it proved difficult to gain access to the appropriate data and contact persons, which forced the project to take another direction.

The updated objective of the thesis work then became to investigate the recent developments in the smart factory scene in general, and identify a number of key areas in which future projects could be initiated. Without a specific manufacturing facility to study, with certain process data and actual machining assets to consider, this objective was proposed as most valuable for Scania.

2.3 Literature study

With the gain of required general knowledge, and the theoretical frame in place, the literature study was conducted. The areas in which literature was studied were Industrie 4.0 (I4.0), IoT, Digital Twin (DT), CPSs, BD, data analytics, Machine Learning (ML), Artificial Neural Network (ANN) and Predictive Maintenance (PdM). The theory chosen for the thesis were mostly collected from research papers in manufacturing journals, publications and frameworks from standardization organizations and teaching material regarding the conceptual and mathematical aspects of ML.

The investigated theoretical areas were then reassessed, with respect to the case at Scania, in order for the analysis to be relevant. This reassessment was conducted in consent with the supervisors at Scania and LTH. One consequence of this was that not all aspects of the theoretical frame of the thesis were of direct relevance for the analysis. They were, however, left in the thesis as they can be useful for reference in future projects.

2.4 Analysis

As the areas of importance for future smart manufacturing within Scania had been established, the analysis was divided in five different interconnected fields, namely *development process*, *information modelling*, *communication structure*, *computational modules* and *collaboration with other companies*. The four former fields are mostly based on previous work and theoretical aspects, while the latter mostly consists of reflections gained during the thesis process.

2.5 Recommended actions

Lastly the thesis work presented the findings in the form of ways for Scania to proceed with future projects and development areas. The thesis also presented some important aspects to consider when proceeding with these project, with respect to the literature and previous research in the field. The different projects were designed to act as standalone projects, but they will of course have some overlapping in between them. Hopefully these projects, which themselves could turn into thesis works, could contribute to the future development of smart manufacturing in general and for Scania in particlular.

3

Theory

This chapter will discuss the relevant theoretical aspects for the thesis, starting with an introduction to Industrie 4.0 (I4.0). Being a complex and rather comprehensive phenomenon, I4.0 will in turn be divided into three key areas, namely Cyber-Physical Systems (CPSs), Big Data (BD) and Advanced Analytics (AA), with heavy emphasis on CPSs. In reality I4.0 consists of an abundance of other areas, which shortly will be mentioned, but these three are of the most importance for this particular thesis.

3.1 Industrie 4.0

3.1.1 Introduction to Industrie 4.0

Industrie 4.0 (I4.0) is a broad term used slightly different across different applications, but there are some often recurring fundamentals used when describing it. McKinsey and Company, amongst others, has written extensive pieces about I4.0 describes I4.0 as being the next big phase in the digitization of the manufacturing sector. The progress leading up to I4.0 has been driven by four separate disruptions: the heavy increase in data volumes, computational power and connectivity; new capabilities in business intelligence and AA; improvements in the transfer between the digital to physical world; and new capabilities in the interaction between human-machine such as Augmented Reality (AR) [Baur and Wee, 2015]. This definition of I4.0 could be compared with one given in a study from the German Academy of Science and Engineering (acatech); defining I4.0 as realtime, high data volumes, multilateral communication and interconnectedness between CPSs and people [Schuh et al., 2017].

The foundation of I4.0 is built upon the advancements of Internet of Things (IoT) through embedding electronics, software, sensors and network connectivity into things [Negri et al., 2017].

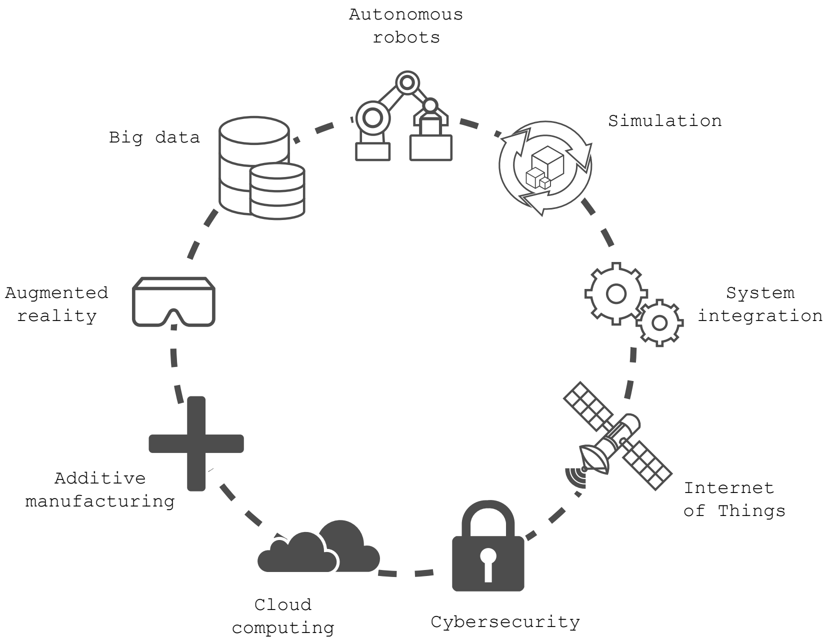

Figure 3.1 Integral areas within Industrie 4.0

3.1.2 Integral areas of Industrie 4.0

I4.0 is a broad concept, with disruptive developing technologies across many different areas, such as *BD*, *AR*, *additive manufacturing*, *cloud computing*, *cybersecurity*, *IoT*, *system integration*, *simulation*, *autonomous* robots [Arm et al., 2018]. These areas are shortly described below.

Industrial IoT and CPSs The concept of Industrial IoT and CPSs originates from using IoT ideas in an industrial setting, by creating a digital representation of the physical object.

Additive manufacturing More commonly known as 3D-printing, additive manufacturing has the potential of truly changing manufacturing methods.

BD The massive trend in increasing data volume and generation speed greatly impacts how future production facilities will operate.

Artificial intelligence The question becomes to find the right algorithm based on every use case. In this field, Predictive Maintenance (PdM) is demanded, as it can be regarded as a function of time.

Robots The development of more advanced collaborative robots ensures safe operations with operators.

Virtual reality Virtual reality may be used for simulation and modeling. AR could prove useful in service routines and maintenance.

From this array of areas, this thesis will discuss development within the industrial setting based on three perspective, as seen in figure 3.2.

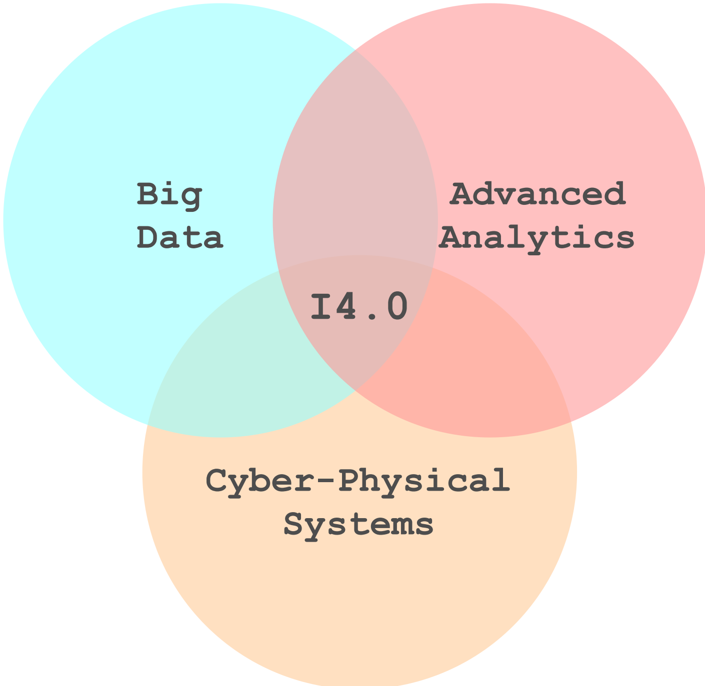

Figure 3.2 Scope of Industrie 4.0 for the thesis

3.1.3 Development stages in Industrie 4.0

acatech describes the process of modernizing a production facility with a model called the ”Industrie 4.0 maturity index”, see figure 3.3, adapted from [Schuh et al., 2017]. This model provides companies with a valuable tool that can help transform their entire organization. The model is based on six different stages of I4.0 maturity, namely *computerization*, *connectivity*, *visibility*, *transparency*, *predictive capacity* and *adaptability*. The first two stages mark the basic requirements for an implementation of I4.0, while the four latter stages correspond to the development process for the next revolution of industry [Schuh et al., 2017]. This perspective on I4.0 development process is shared across a variety of different sources. The I4.0 maturity index bears close resemblance to e.g. the 5C structure used to describe the value creating levels of a CPS [Monostori et al., 2016], which will be discussed in more depth in section 3.2.3.

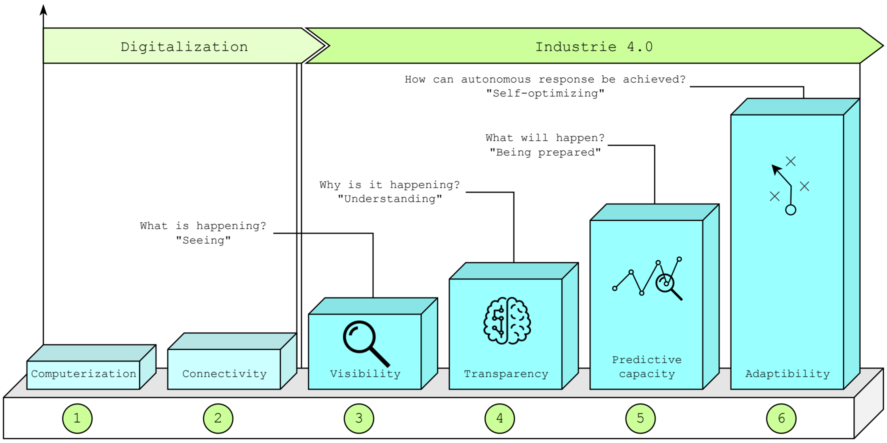

Figure 3.3 Industrie 4.0 maturity index

3.1.4 Business value associated with Industrie 4.0

The impact of I4.0 have been widely discussed and while it proves difficult to give a correct estimation of the financial gain, there is a consensual agreement that it will be of significant nature. Boston Consulting Group (BCG) suggests that I4.0 will drive a productivity gain of 5 to 8 % on total manufacturing costs over the coming ten years, at a total of €90 000 million to €150 000 million, in Germany alone [Rüßmann et al., 2015].

3.1.5 Smart manufacturing

I4.0 is originally a German term, used first and foremost to strengthen Germany's position at the global scene. A more generic term is *Smart Manufacturing*, which is clearly defined by International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) as *"Manufacturing that improves its performance aspects with integrated and intelligent use of processes and resources in cyber, physical and human spheres to create and deliver products and services, which also collaborates with other domains within an enterprise's value chains"*. I4.0 will, however, still be used in the thesis.

3.1.6 Reference Architecture Model Industrie 4.0

3.1.6.1 Introduction to Reference Architecture Model Industrie 4.0 Reference Architecture Model Industrie 4.0 (RAMI 4.0) is a framework for systematically classifying and refining important elements of I4.0 technology. RAMI 4.0 consists of a 3D coordinate system in which assets as well as complex inter-relationships could be broken down and classified [Adolphs et al., 2016], see figure 3.4. This will be an important step towards a common "language" when discussing and developing organizations and business models as well as identifying overlapping systems within the manufacturing setting. Both Open Platform Communications Unified Architecture (OPC UA) and Automation Markup Language (AutomationML) can be classified as standards on the RAMI 4.0, within the communication and information layers, respectively [OPC Foundation, 2017].

RAMI 4.0 is composed of three different perspectives, depicted by axes, namely *layers*, *life cycle & value stream* and *hierarchy levels*. These concepts will be presented below.

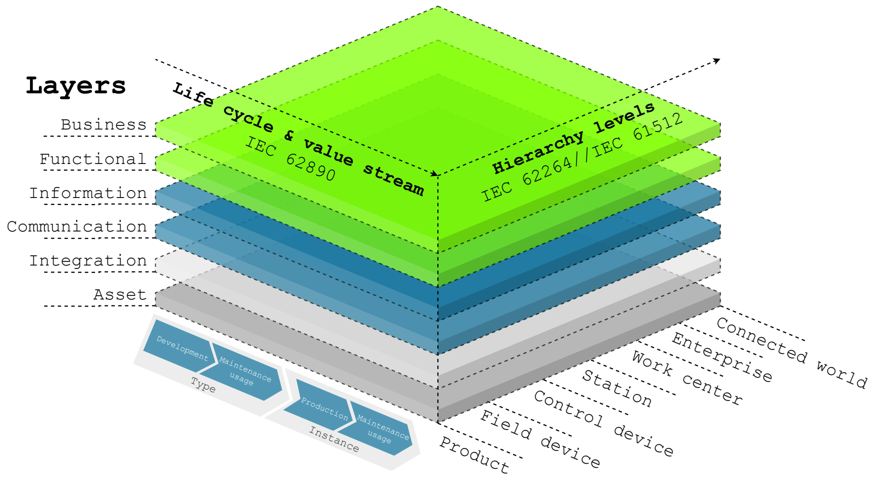

Figure 3.4 Reference Architecture Model Industrie 4.0

3.1.6.2 Layers The first perspective of RAMI 4.0 is set on the vertical axis of the coordinate space. Layers are used to represent the splitting of complex projects from an IT viewpoint into manageable parts.

Business layer is the highest layer and includes the mapping of business models and rules into the architecture. It also contains legal and regulatory frameworks and the results of the process as a whole. This would, however, typically not include concrete systems such as Enterprise Resource Planning (ERP), which are located inside the functional layer.

Functional layer includes a formal description of the different functions used in the solution space. The functional layer enables a run time and modeling environment for services supporting the business layer as well as a run time environment for applications with technical functionality. It is within the functional layer that methods for analytics and data science would be categorized.

Information layer handles the modeling and formatting of data and information which are communicated within the system. This could include how different assets within the system are described and identified as well as how data is structured and labelled.

Communication layer represents the communicative protocols and services between the Integration layer and the Information layer. The communication layer could further be broken down using the classic Open Systems Interconnection (OSI) model, i.e. (1) *physical*, (2) *data link*, (3) *network*, (4) *transport*, (5) *session*, (6) *presentation* and (7) *application*. RAMI 4.0 suggests OPC UA as the communication protocol for the space occupied between the session layer through application layers, from the product level through work center levels within the instance in the life cycle, which in practice would mean the complete automation lines in a factory.

Integration layer serves as the link between the physical and digital world, and provides information generated by the asset layer in a form which can be processed by a computer. It could be described as the digitalization of the assets in the real world. It is within this layer that elements connected with IT is handled, such as implementation of sensors, actuator, Radio-frequency identification (RFID) readers and Human Machine Interfaces (HMIs). It is also within the integration layer where the data gathered from the processes would be classified.

Asset layer represents the physical and virtual things within the architecture, i.e. the "reality". This includes all physical components such as machines, manufactured products, human operators, and software components such as standardized software, documentation and even project plans and ideas. The physical things are then connected to the digital world though the integration layer, either actively or passively through the use of technologies like Quick Response (QR) codes.

3.1.6.3 Life cycle & value stream The product life cycle spans from the initial idea of a product, through production and sales, all the way to service and maintenance of the product. The idea is to uniformly collect the data from the entire process of the product. This axis is based on IEC 62890.

Type is created with the initial idea of an asset or concept, e.g. when a new product is being developed. This enables gathering of all relevant data regarding the type, including design schematics, manuals and other engineering documentation, physical models, just to name a few.

Instance is the realization of the types, as it moves into production. Each manufactured product represents an instance of the type and could receive a globally unique identification number. For the customers, the products are initially again depicted as a type until installed within a particular system. The transition between Type and Instance could occur multiple times.

3.1.6.4 Hierarchy levels The Hierarchy levels of RAMI 4.0 originates from the automation pyramid proposed in the widely adopted ISA-95 standard (IEC 62264), with the addition of the internet as a whole, i.e. *connected world*, and the actual product being manufactured, i.e. *product*.

Connected world This level enables the presence of multiple connected factories within the same organization architecture which need to communicate.

Enterprise This level includes Enterprise Resource Planning (ERP), with time scale in months-days.

Work center This level includes Manufacturing Execution Systems (MES), with time scale in hours.

Station This level includes Supervisory Control And Data Acquisition (SCADA), with time scale in minutes.

Control device This level includes Programmable Logic Controller (PLC), with time scale in seconds.

Field device This level includes sensors and signals, with time scale in μs/ms.

Product To include the product in the architecture enables for increased transparency within the manufacturing process by relating process and production data directly to the product itself.

3.1.6.5 Connection to ISA-95 ISA-95 is an international standard that treats how enterprise systems should be integrated with manufacturing and control systems [Monostori et al., 2016]. Part 3 of the standard is particularly interesting, suggesting how the exchange of information between the lower level process control and the upper level business planning and logistics should be executed [Johnsson, 2003].

3.1.6.6 Physical and information world The physical world is fully encapsulated within the asset layer and all other layers include information and other immaterial objects only. Every piece of digital information and immaterial object need to have a physical carrier somewhere within the physical world. This medium could be on a digital mass storage, a piece of paper or the knowledge of a human being. Information not present on some kind of medium within the physical world is therefore not known to the information world and not considered in RAMI 4.0. If the physical carrier is removed from the model, the information that it holds will also be deleted. In order to avoid information to disappear from the system, multiple equivalent object images can be stored on different physical containers. Even though these images can appear across different physical mediums, the information entity should be considered as singular [VDI/VDE, 2016].

3.1.7 Industrie 4.0 component and Asset Administration Shell

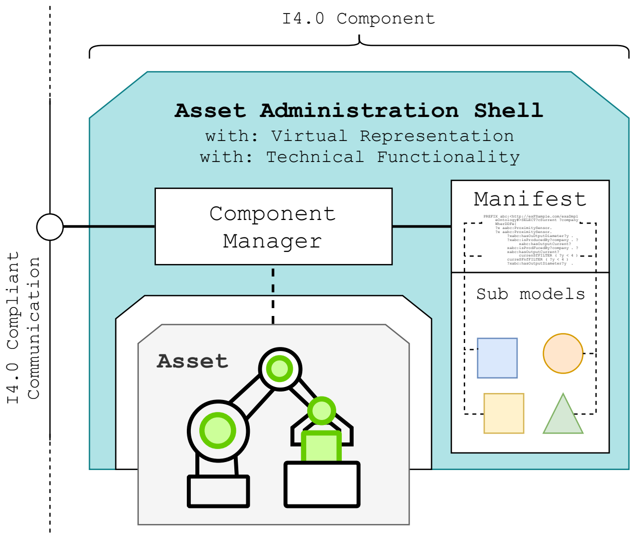

Figure 3.5 Industrie 4.0 component and the Asset Administration Shell

The I4.0 component is defined as a ”globally uniquely identifiable participant with communication capability consisting of administration shell and asset within an I4.0 system [VDI/VDE, 2017], see figure 3.5. This means that the I4.0 component could be any asset connected to a I4.0 type network. For this to be made possible, a shell has been developed, called the Asset Administration Shell (AAS). AAS is a logical set of information and an administration interface surrounding the asset and could be defined as ”the virtual digital and active representation of an I4.0 component in an I4.0 System” [Wagner et al., 2017]. In practice, the AAS collects all relevant data through the component's life cycle in a digital, secure container carried by the component [Hoffmeister, 2015].

With regards to RAMI 4.0, the AAS could be viewed as the mapping of the RAMI 4.0 to a real thing. An asset can in turn consist of multiple other assets, i.e. the components in a machine can both have individual AASs as well as one AAS capturing the entire machine.

AAS includes a Digital Factory (DF) *header*, containing information regarding identification and designation of the asset, and a DF *body*, containing information regarding the features and capabilities of the asset, as well as the disposition, i.e. data element values [Adolphs et al., 2016]. Both the *header* and the *body* in turn include a *manifest*, containing identification of both the asset and the AAS in the header and models and submodels within the body, as seen in figure 3.6. Each submodel has hierarchically organized properties, which describe the data and functions of the model, which could be realized by AutomationML. The body also includes a component manager, which practically could be realized by OPC UA. Every AAS consists of exactly one component (or resource) manager [VDI/VDE, 2016].

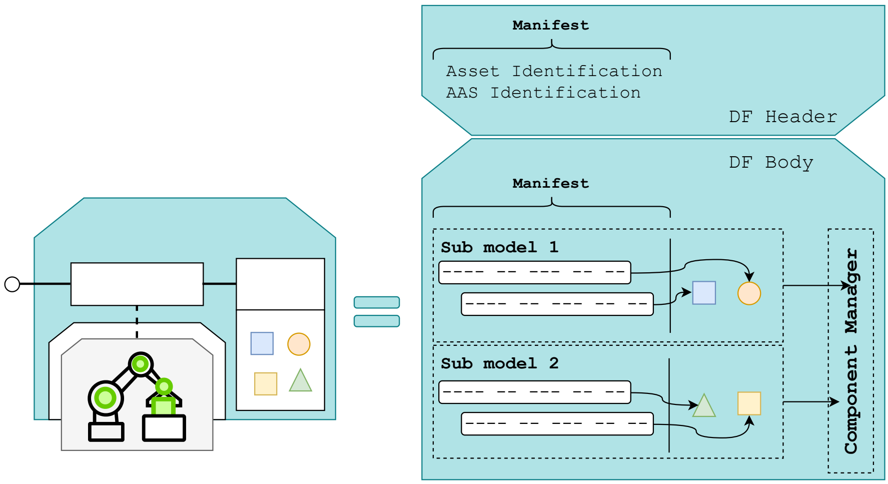

Figure 3.6 Expanded view of the Asset Administration Shell

AAS bears close resemblance with the concept of a Digital Twin (DT). DT was first mentioned by NASA in the aerospace sector as "an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin" [Negri et al., 2017]. In this original sense, the DT is used solely as a tool for simulation, with emphasis on the product itself. More recently, the DT concept has increasingly been used to simulate models of complete manufacturing plants. For this thesis, however, AAS will consistently be used when referring to the digital representation of an asset.

3.1.7.1 Presentation and communication One important question when regarding I4.0 systems is "Which information regarding an asset is known to the digital IT-system?". This question motivates the use for these different ***presentation classes*** for assets in the information world [VDI/VDE, 2016]:

unknown assets, which are not known to the information world.

anonymously known assets, which are known to exist somewhere in the system, but without any individual information. These could be physical resources within a container in a production plant, where the number of resources is known without knowing any specific details of individual objects, apart from general information. When one of these assets is installed in the plant, or put in production, they can become individually known, but only as long as they are integrated in any process of the plant.

individually known assets have a unique name which is known throughout the system.

managed entities are assets which have assigned objects within the information world, for asset administration. This administrative functionality of the asset could enable asset tracking, quality control and operational control of the asset.

Furthermore, the Communication and Presentation (CP) classification relies on the ***communication abilities*** of the the assets, which is classified by:

not able to communicate, could be assets without information carrier functionality, or assets with this functionality but without communicative interfaces.

passive communication ability implies that the asset can communicate with the network, but not standalone. It could be that its information needs to be extracted through a QR code of RFID tag.

active communication ability means that the asset actively can participate in the network's communication. Upon initiation, the asset identifies itself and registers as a participant in the network traffic.

I4.0 compliant communication ability implies that the asset provide all necessary abilities to participate in an I4.0 system. This includes globally unique identification, support for generic I4.0 standard services, adequate security and data protection, offering required realtime ability and support for I4.0 standardized semantics.

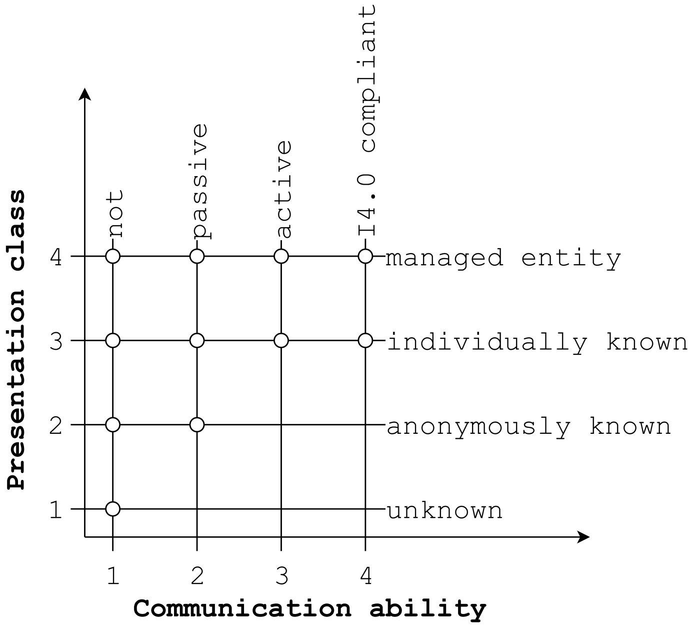

Figure 3.7 Communication and Presentation classification for the Industrie 4.0 component

These two independent categories make up the so called CP classification, as seen in figure 3.7, which can be used in I4.0 applications. If an AAS and an asset is connected with a digital communication system, they together make up an I4.0 component. A prerequisite for an I4.0 component is that it is of CP24, CP34 or CP44, meaning that it needs to be administered through the I4.0 system and at least needs to be of passively communicative ability.

3.2 Cyber-Physical Systems

3.2.1 Introduction to Cyber-Physical Systems

One core aspect of preparing a factory for I4.0 is to create a digital representation of the systems involved. Many believe that the goal of implementing an I4.0 type operation could be achieved with the use of CPSs [Bagheri et al., 2015].

Cyber-Physical Systems (CPSs) is a name given to systems of entities which are computationally collaborating, with strong connectivity to the surrounding physical world. The core of CPSs consists of transformative technologies for managing these interconnected systems between their physical assets and computational capabilities [Lee et al., 2015]. These systems are simultaneously providing and using data-accessing and data-processing services available on the IoT. CPSs are engineered around a computing and communicating core, thus integrating monitoring, control and coordination between the physical systems. When designing and analyzing CPSs, it is not the understanding of either the physical nor the digital system that is of importance, but rather the interaction and interconnectivity between them [Monostori et al., 2016].

As a sub-category to CPSs, we find Cyber-Physical Production Systems (CPPSs): Autonomous and cooperative elements and sub-systems that are connected across all levels of production from machine processes up to production and logistics networks. They are *intelligent* (able to acquire and analyze information from surroundings), *connected* (connected to other elements of the systems for cooperation and collaboration) and *responsive* (towards external and internal changes) [Monostori et al., 2016].

3.2.2 Technological advancements with Cyber-Physical Systems

There will be great technological advancements with introducing CPSs into the manufacturing workflow. CPSs open up the ability to continuously monitor the status of the assets, increasing visibility and transparency of the manufacturing process, while also enabling for remote control of the equipment in the plant.

Furthermore, this process could be automated to automatically optimize different operations based on readings from sensors in order for preemptive and predictive maintenance to be executed.

Additionally, CPSs could make remote diagnosis of equipment possible while also providing field technicians with additional aid and support for better results and efficiency, through visual and audio information. The machines could also automatically order spare parts to shorten the service maintenance time.

Lastly, the information and insight gained from the operations data could be sold to other stakeholders, e.g. machine manufacturers, to further increase the knowledge of how the machines operate in an actual manufacturing plant, which in turn could drive the technology forward for the industry in the future [Herterich et al., 2015].

3.2.3 5-level Cyber-Physical System structure

CPSs aims to solve two particular problems, namely to (1) act as an advanced connection between the data acquisition from the physical world and information feedback from the digital world and (2) ensure intelligent data management, analytics and computational capacity that makes up the cyber space [Lee et al., 2015]. As these problems seem somewhat ambiguous, there is a need for more substantial structure as how to define and implement CPSs for actual use.

One commonly proposed structure for implementing CPSs would be the so called 5C structure, introduced in [Lee et al., 2015]. This structure consists of 5 levels in a sequential workflow manner and illustrates how to construct a CPS from the initial data acquisition through analytics to final value creation, see figure 3.8. The structure could be viewed as a stepwise guide for constructing the CPS to work in a manufacturing setting.

The first layer in this structure represents physical space, second to fourth represent pure cyber space while level five realizes feedback from cyber to physical space, in an autonomous fashion [Monostori et al., 2016]. A detailed description of the layers is as follows:

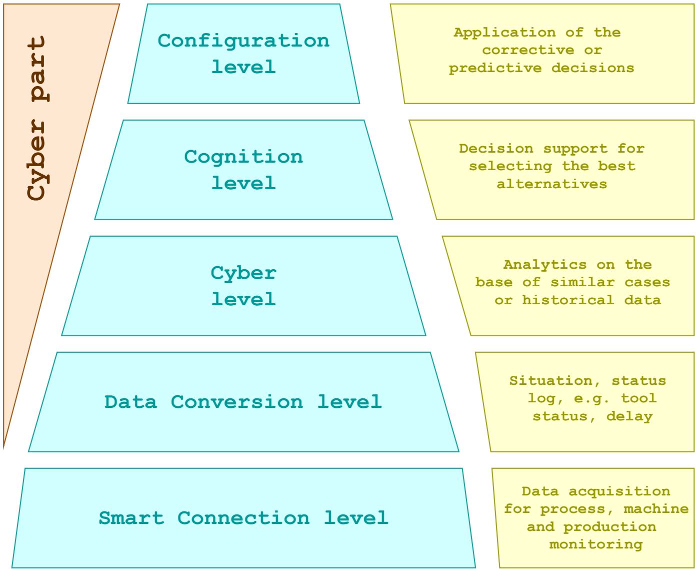

Figure 3.8 5C structure for implementation of Cyber-Physical Systems

Smart connection layer The first step in deploying a CPS is through acquiring accurate and reliable data from machines and their components. Data may be collected directly from sensors or obtained from controller and enterprise systems such as ERP, MES and PLC. Selecting sensors, data sources and transfer protocols could prove crucial for the CPS's performance in higher layers.

Data-to-Information conversion layer Relevant information has to be extracted from the data. Currently, several tools and methodologies are available for the data conversion layer, with focus on prognosis, health management, estimation of remaining useful life, machine condition, etc.

Cyber layer The middle layer acts as central information and analytics hub in the 5C structure. Information is being gathered from the connected assets to form a network. Individual machines can put their performance in relation to others within the network as well as to historic data to predict future behavior.

Cognition layer The previous layers provide a thorough understanding on the monitored system as a whole, as well as for individual machines. This layer puts emphasis on presentation of the acquired knowledge through infographics to experts to support decision making.

Configuration layer Lastly, the configuration layer acts as feedback from the cyber space to the physical space and functions as high level supervisory control. This layer automates the actions based on decisions from the cognition layer to the monitored system.

In this case the cyber level is of great importance, as is serves as the link between the lower level data acquisition and the upper level cognition features [Bagheri et al., 2015]. It is within the cyber layer that AA models could be constructed, as discussed later in the thesis.

3.2.4 Communication structure for Cyber-Physical Systems

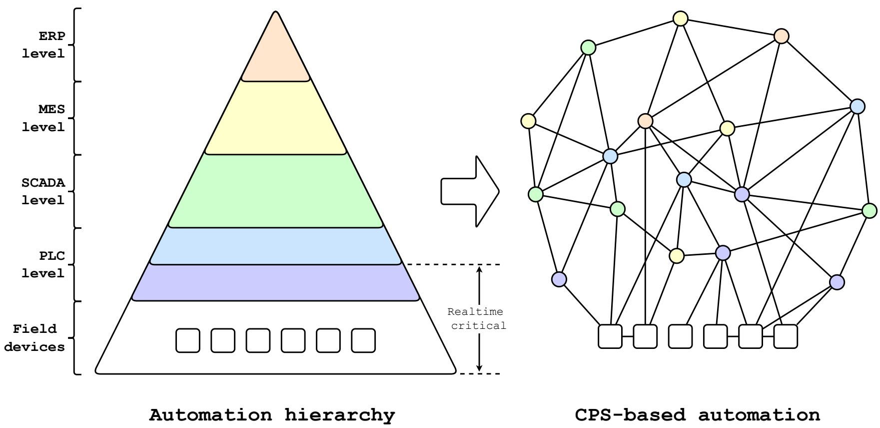

Figure 3.9 Decentralized communication for Cyber-Physical Systems

CPPSs will partly change the classic automation pyramid into a more decentralized communication structure [Monostori, 2014], see figure 3.9. The systems closer to the physical processes will generally remain the same for optimal performance in time critical tasks such as process control. The higher level systems will, however, follow a more decentralized way of functioning.

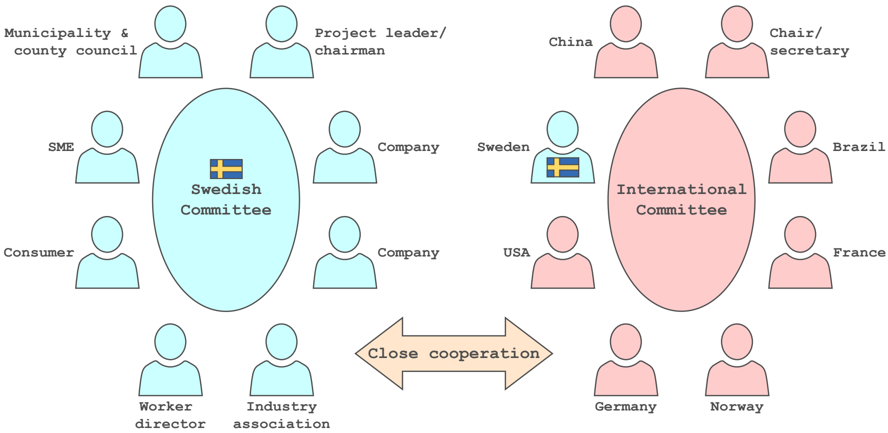

Figure 3.10 National and international standardization committees

3.2.5 Industry standards within Cyber-Physical Systems

One key aspect of achieving an I4.0 type operation is to adopt the use of standards and common frameworks for use across different machines and machining manufacturers. Two of the most influential organizations are ISO and IEC, which are appointed by World Trade Organization (WTO). Each of the world's 195 countries can, in turn, have their respective national mirror committees, as illustrated in figure 3.10[1].

This means that the standardization work is carried out simultaneously on national and international levels. On the national level, the different participants include companies and organizations representing many different branches. The national committees can then send experts as representatives in the international committees, to develop international standards. Lastly, in order for a standard to be approved, it needs to be voted on by the participating nations, where each country gets one vote.

Some useful frameworks and protocols will be presented and briefly described in the following sections.

3.2.5.1 Open Platform Communications Unified Architecture (OPC UA) is a platform-independent machine-to-machine communication protocol for industrial automation [Zezulka et al., 2018]. OPC UA is a recognized international standard (IEC 62541) and further defines how information is exchanged between the systems involved and handles data management and communication management [Monostori et al., 2016].

[1] The technical committees that, inter alia, mirror ISO in Sweden are arranged by Svenska Institutet för Standarder (SIS) – the Swedish institute for standardization.

3.2.5.2 Automation Markup Language (AutomationML) is an upcoming open international standard (IEC 62714) for description of manufacturing plants and plant components. AutomationML models assets' topology, interfaces and relations to others, geometry, kinematics and even logic and behavior [Monostori et al., 2016]. AutomationML is a neutral data format based on XML, and incorporates several different standards, Computer Aided Engineering Exchange (CAEX) (IEC 62424) as top-level format and for plant topology, COLLAborative Design Activity (COLLADA) for geometry, kinematics and motion planning, PLCopen XML (IEC 61131-10) for behavior and sequence descriptions and Mathematical Markup Language (MathML) for formulas.

3.2.5.3 universal machine tool interface (umati) is an upcoming common interface for machine tools. umati's core functionality is standardized semantics, embedded in the information model based on OPC UA. This standard will enable machine tools to easily, securely and seamlessly be connected to any customer's varying IT ecosystems. Moreover, universal machine tool interface (umati) has support for specific extensions for customers and manufacturers of machine tools.

3.2.5.4 Equipment Behavior Catalogues (EBC) is part of an upcoming international standard (ISO 16400), which addresses behaviors and process information for virtual productions systems, which can be used for simulation.

3.3 Big Data

3.3.1 Introduction to Big Data

Recent disruptive development resulting in wider availability and affordability of sensors, data acquisition systems and computer networks drive factories towards implementing high-tech methodologies. Consequently, the ever growing use of sensors and networked machines has resulted in the continuous generation of high volume data which is known as Big Data (BD) [Lee et al., 2015].

In general, BD can be described as a large amount of structured, semi-structured and unstructured data, which cannot be collected, stored, managed, shared, analyzed and computed using traditional data tools within a tolerable amount of time [Qi and Tao, 2018]. With a rapid increase in data-producing sources, the focus shifts towards which information gain and insight the data may provide, rather than the data itself.

3.3.2 5V's of Big Data

A common way of describing BD would be with the so called 5V's of BD, see figure 3.11. The different V's stand for *Volume*, *Velocity*, *Variety*, *Veracity* and *Value* and are explained as follows.

Volume refers to the extensive amounts of data which are created and stored every moment [Iqbal et al., 2017].

Velocity refers to the speed at which data are generated, aggregated as well as moved around [Iqbal et al., 2017] and that the data processing requires high timeliness [Qi and Tao, 2018].

Variety refers to the various types of data that are collected. This includes data of different sizes and applications that are *unstructured* (e.g. digits, symbols and tables), *semi-structured* (e.g. trees, graphs and XML documents) or *structured* (e.g. logs, audio, imagery, video and documents) [Qi and Tao, 2018]. The processing of this data proves to be technically difficult using traditional methods as the data cannot be categorized in traditional relational databases [Iqbal et al., 2017].

Veracity refers to the messiness or trustworthiness of gathered data. Due to the cheer amount and velocity of data it often contain some extent of noise [Iqbal et al., 2017].

Value refers to providing meaningful and relevant insight into BD. This includes recognizing patterns leading to information gain.

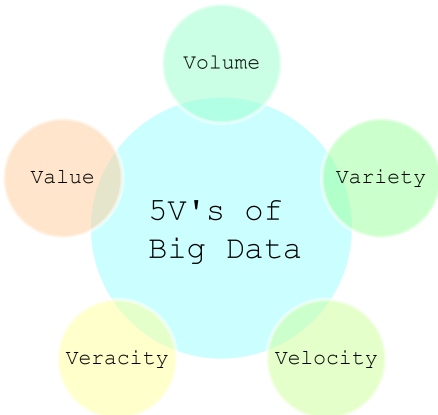

Figure 3.11 5V's of Big Data

3.3.3 Big Data in an industrial setting

As mentioned previously, smart factories operate using a large number of sensors and information driven technology generating large amounts of data, creating the need for new technology and algorithms for processing the data. With reliability and safety being high priority concerns in modern factories, the complexity of the automated and flexible manufacturing plants challenges these areas.

Processing industrial BD involves data formatting, dimensionality reduction, hidden pattern identification, performance evaluation and prediction [Yan et al., 2017].

The sources of industrial BD could include the following

- Design data from the products or plant components
- Machining operations data form controllers
- Personnel behavior records
- Cost information from manufacturing process
- Logistics information
- Environmental data, such as temperature, humidity, noise
- Fault detection and system status
- Product quality data
- Customer information

3.4 Advanced Analytics

3.4.1 Introduction to Advanced Analytics

Advanced Analytics (AA) are a set of methods for extracting valuable information from big amounts of data, closely related to statistical analysis. These methods are often autonomous or semi-autonomous and are often highly sophisticated. The speed of which AA algoritms are developed is ever increasing, and with the massive expansion of BD and computational power at hand, the methods will continue to become even more sophisticated.

At a shallow level, Machine Learning (ML) algorithms can be divided into two different categories, unsupervised or supervised learning, depending on what kind of experience they are given during the learning process. There are many different use cases and methodologies associated with AA, and this section will outline some differences and strengths with some of these approaches.

3.4.2 Unsupervised learning

Unsupervised learning algorithms often experience a dataset consisting of many features and is then set to learn relevant properties of the structure of the data. This could include clustering, i.e. dividing the data into segments of similar properties. In short, the unsupervised learning process aims to map the entire probability distribution that generated the given dataset [Goodfellow et al., 2016].

3.4.2.1 Principal Component Analysis As much of the data in modern factories are of high dimensionality, the ML problems often become unnecessarily complex and difficult to solve. This phenomenon is called the "curse of dimensionality". Oftentimes, it is crucial to initially reduce the dimensionality of the data set. This can be done with Principal Component Analysis (PCA), which uses Singular Value Decomposition (SVD) to find the dominating features which most accurately can represent a high dimensional data set. PCA could be described as the orthogonal representation of data, onto a lower dimensional linear space [Bishop, 2006].

3.4.2.2 k-means clustering Another learning algorithm for unsupervised problems is k-means clustering. Essentially, k-means clustering divides the data set into k different clusters, where the examples in some regard are similar to another. This is done by randomly initiating k centroids $\{\mu^{(1)}, \ldots, \mu^{(k)}\}$ and then alternating between two different steps until convergence. First, each data point is assigned a cluster depending on its nearest centroid. Secondly, the centroid is updated to the mean of all data points belonging to the cluster [Goodfellow et al., 2016].

3.4.3 Supervised learning

Supervised learning, on the other hand, experience a dataset consisting of different features, where each instance is associated with a label. The learning process then attempts to make correct predictions of the output, based on a given input. Supervised solutions are generally preferable over unsupervised methods, if the required data for the task is given. [Paolanti et al., 2018].

3.4.3.1 Linear regression One of the most simple tools for ML is linear regression, which tries to find an affine function that maps an input to an output, i.e. $\hat{y} = \boldsymbol{w}^T \boldsymbol{x} + b$. The weights in the vector $\boldsymbol{w}$ are then chosen, in order to minimize the mean square error $\text{MSE} = \frac{1}{m}||\hat{y} - y||_2^2$, using the gradient of MSE with respect to w, i.e. moving in the negative direction of the gradient until $\nabla_w MSE = 0$. This is not a particularly sophisticated technique, but it provides a good understanding on how ML works in practice [Goodfellow et al., 2016].

3.4.3.2 Support Vector Machine One supervised learning technique that gained a lot of traction in the early 90's is Support Vector Machine (SVM). SVMs are driven by an affine function $\boldsymbol{w}^T \boldsymbol{x} + b$ and predict a class depending on if the function is positive or negative. The support vectors together form a margin, which separates the classes as clearly as positive. This makes SVM very stable to outliers, as the support vectors are only defined at the border of the classes.

By default, SVM can only classify linearly separable data representations. One could, however, run the data set through some nonlinear kernel, which could make the data set linearly separable [Goodfellow et al., 2016].

3.4.3.3 Artificial Neural Network Currently, one of the more popular categories of ML techniques are Artificial Neural Network (ANN). These techniques vaguely mimic the biological brains ability to learns and they seem to perform really well in highly nonlinear applications.

Multilayer Perceptrons Deep feedforward networks, or Multilayer Perceptrons (MLPs), have the goal of approximating some function f^* which maps some input $\boldsymbol{x}$ to a category y. The mapping becomes $y = f^*(\boldsymbol{x};\boldsymbol{\theta})$, and the neural network will learn the parameters $\boldsymbol{\theta}$ during its training phase.

Fully connected MLPs means that the input of each ”neuron” consists of a linear combination of every neuron in the previous layer, with some weight w_i and a bias term b_i defining the relation. Depending on the activations of the previous layer, a neuron can be active or non-active (or something in between). The basic structure of MLPs can be observed in figure 3.12.

Most importantly for the learning process for the MLPs is the concept of back-propagation. This is where the outputted estimate $\hat{y}$ is compared to the actual label y, and each of the weights and biases are ”tuned” to better fit the data. Back-propagation works via the concept of stochastic gradient decent, where each parameter is tuned depending on the partial derivative for the loss function, corresponding to the well known chain-rule from calculus.

The algorithm then becomes, first feed in some input to the network, then compare the estimated results to the ground truth and lastly, let the network tune each parameter through back-propagation. This way of letting the network automatically craft ”features” for classification has proven to be very successful in an array of applications, where manually engineered features are simply not good enough as a results of our limited ability to correctly model the reality [Goodfellow et al., 2016].

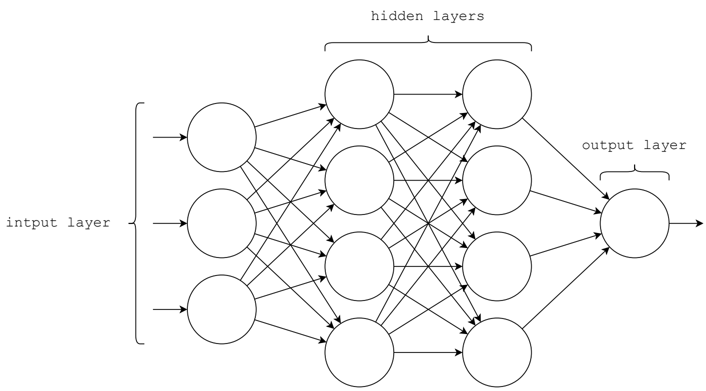

Figure 3.12 Structure of a typical Multilayer Perceptron

Convolutional Neural Networks Another variant of neural networks are Convolutional Neural Networks (CNN). The so called convolutional layers include a number of filters, which are convolved over the entire example. This method is especially popular for image recognition tasks. Each layer may then pick up different patterns in the data. For an image, the first layer might pick up on edges, the second could combine multiple edges into textures and for each layer the filters become increasingly sophisticated [Bishop, 2006].

Recurrent Neural Networks and Long Short-Term Memory The types of neural networks mentioned in previous sections are neither time-dependent nor sequential. For these tasks, the use of Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) could prove useful. These networks are specialized for processing sequences of values $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$. They are based on the idea of parameter sharing across different parts of the network, and can "remember" occurrences back in time to make more accurate predictions about the future.

3.4.4 Adaptive clustering for self aware machines

One proposed method for introducing AA into the CPS environment would be through adaptive clustering [Bagheri et al., 2015]. This method takes place within the cyber level of the 5C structure, and the system autonomously summarizes, accumulates and learns based on the collected data from multiple machine instances. The methodology for the autonomous ML and knowledge extraction is divided into two main steps, namely (1) *similarity based clustering* for identifying machine condition and working regimes and (2) *prognosis of machine health* under complex and multi-regime conditions, see figure 3.13.

3.4.4.1 Similarity based clustering Based on the existing similarities within the machine fleet, either due to spatial (e.g. similar tasks or environment) or temporal (e.g. similar service or working times) reasons, machine clusters based on performance or health condition could be built into a knowledge base.

Using unsupervised learning algorithms, such as Self-Organizing Map (SOM) or Gaussian Mixture Model (GMM), different clusters can be formed, based on working regimes and machine condition. The methodology is based on an online update mechanism which takes in different parameters from the system and tries to identify if the state matches an existing cluster using multidimensional distance measurements. This could either result in the findings of a similar cluster, in which case the machine in question is labeled as belonging to the identified cluster (and the cluster criterion updated), or no similar clusters found, in which case the state of the machine will be stored for later. If enough samples of similar unidentified machine states appear, a new cluster will automatically be created and the corresponding machine states added to it.

The growing cluster will over time become a knowledge base in the cyber level, which later on could be used for identifying individual machine health [Bagheri et al., 2015].

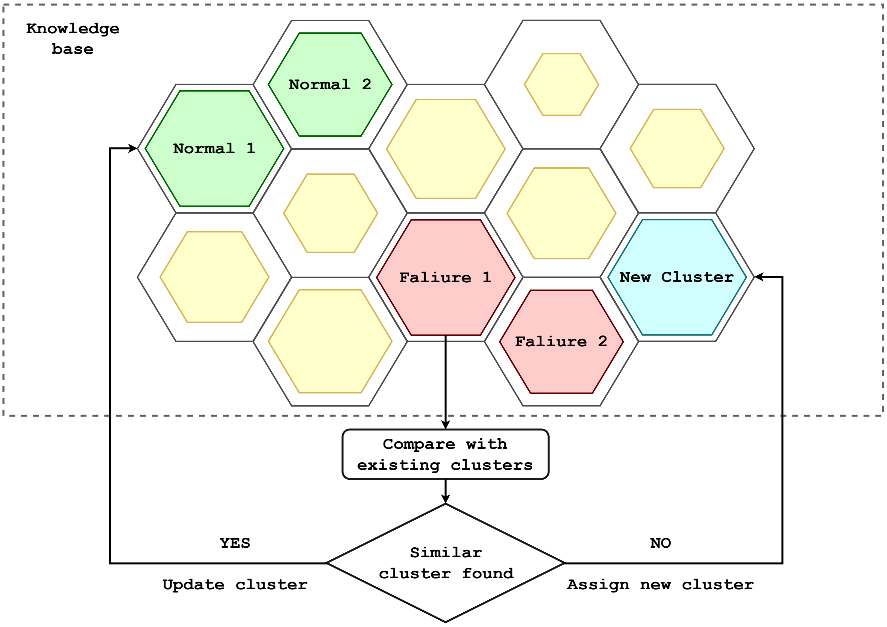

Figure 3.13 Adaptive clustering for self aware machines

3.4.4.2 Machine health prognosis The next step will be to predict the Remaining Useful Life (RUL) of the machine, using its utilization history and measurement data. One advantage with the proposed technique is to regard the stress vs. life aspect, i.e. to include different stress factors in addition to the operating time for the machine degradation. The workflow is divided into four parts, where the first two steps extract and accumulate the stress vs. life relationship that will be used for health prognosis and the last two steps use the knowledge base for predicting the RUL for a specific machine [Bagheri et al., 2015].

4

Case Study

This chapter will present relevant information regarding the case which will be used to connect the collected theoretical aspects to practice. The material used in this chapter has been provided by Scania.

4.1 Scania

Scania is a world leading provider of transportation solutions, mostly known for its production of buses and trucks. Scania is also know for its heavy emphasis on tailor products and services offering [Scania, 2019b].

Currently, Scania employs around 52 100 people in about 100 countries, providing Scania with a global scope and ability to satisfy the customer's needs, independently of the geographical location. The research and development is mostly focused to Sweden, with branches in Brazil and India. Scania maintains a high level of investment back into R&D to strengthen its product line and manufacturing routines, with investments equivalent to over 7000 MSEK in 2018 alone, which could be compared to its net sales of 137 000 MSEK in the same period [Scania, 2019b].

By combining heavy trucks, buses, engines and product related services, Scania can offer a wide range of cost-efficient and highly modular solutions. It has been estimated that Scania currently manufactures 1.2 *exactly similar* vehicles per year world wide, as a results of the available customization options for its customers [Scania, 2019a].

4.1.1 Vision and ambition

Scania's vision is to work towards a more predictable smart factory operation. This vision has been divided into four different subcategories, namely

- *Smart logistics*
- *Smart machining*
- *Smart assembly*
- *Smart maintenance*

These categories include different milestones for what the smart factory tries to accomplish.

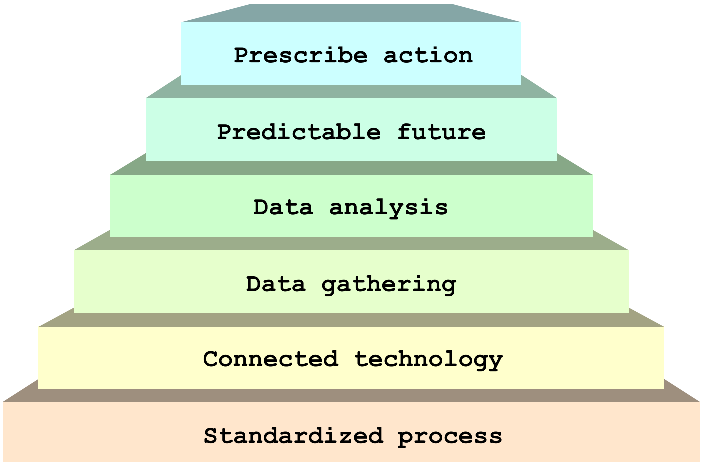

Figure 4.1 The smart factory pyramid at Scania

4.1.2 Scania IT ecosystem

Scania's digitalization process could be viewed in the digitalization pyramid, pictured in figure 4.1. This process could then be connected to the Scania IT ecosystem in figure 4.2, which shortly will be mentioned in the following sections.

4.1.2.1 Connect and gather The bottom part of the pyramid focuses on connecting all the ”things” in the industry, which is the most basic part of the ecosystem. At Scania, this would be done in the Production Service Bus and Production and logistics equipment and devices, as seen in figure 4.2. Some of the ideas from this layer in the pyramid are infrastructures such as LAN, Wifi, 5G, as well as generic services for security and device management.

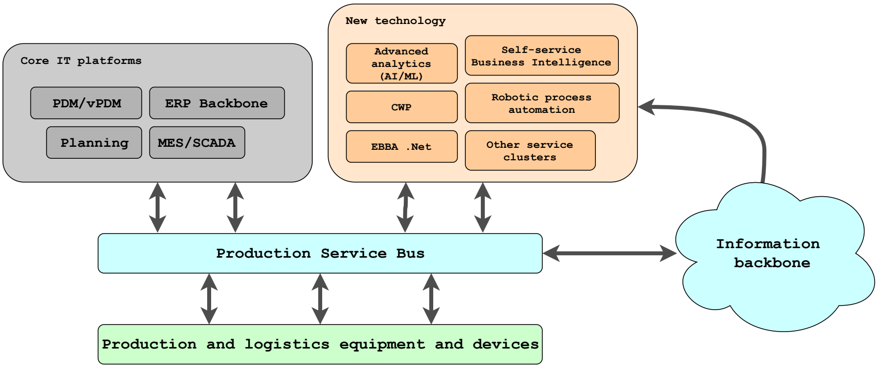

Figure 4.2 Scania IT ecosystem

4.1.2.2 Gather and analyze The next step towards digitalization within Scania is within the gather and analyze layers. This can be represented by the Information backbone in figure 4.2. Within these layers resides advanced analytics methods for unstructured data in a datalake, and self service and business intelligence such as a common data warehouse for enterprise wide and cross process analysis.

4.1.2.3 Predict and prescribe Lastly, the most sophisticated steps in the pyramid are the predictable and prescriptible methods. This is where the new technology should reside, such as AA for material prediction and deliveries, intelligent takting[1] and resource optimization. Additionally, service clusters for MES and SCADA systems as well as a DT with AR capabilities could be included as new technologies.

4.2 PTC

PTC is a global software company with headquarters in Boston, Massachusetts, whose products enable companies to design, manufacture, and service assets (or as they are called at PTC – ”things”) for a smart, connected world. PTC's software specializes in converging the physical and digital worlds, through IoT, AR, 3D printing, digital twin and I4.0. PTC employs around 6000 people in 30 countries, with over 300 000 developers in its global ecosystem [PTC, 2019].

Amongst PTC's array of software and services, three main applications have been of interest for this thesis, namely *ThingWorx*, *Vuforia Studio* and *ColdLight*.

4.2.1 ThingWorx

ThingWorx is PTC's main IoT platform, which enables for modeling of *Things*, or assets, to create a digital representation of an enterprise. This is done with a smart

[1] Takt time, or cycle time, is the average time between the start of production of a unit and the start of production of the next unit.

use of so called *Thing Shapes* and *Thing Templates*. Thing Shapes are at the top of the inheritance structure and acts a base definition component. A Thing Shape adds capabilities to Thing Templates and a Template can inherit multiple Shapes. Thing Templates, on the other hand, model a set of similar objects. A Thing instance can inherit one Thing Template, and add optional functionality if needed. Things can then have *properties*, *services*, *events* and *subscriptions* [PTC, 2019].

Built into ThingWorx are also support for cybersecurity and role assignment, for a secure digital representation of the physical world.

4.2.2 Vuforia Studio

Vuforia Studio is the platform for AR modeling and connection to properties in ThingWorx. Vuforia Studio can present realtime information from Things, and add animations and step-by-step instructions for maintenance and educational purposes.

4.2.3 ColdLight

ColdLight is an application built into ThingWorx analytics, and it is the main data analytics engine in PTC's software. ColdLight has support for many different learning algorithms and methodologies and acts as a powerful suite for multiple analytics problems.

5

Analysis

This chapter will conclude the theoretical advancements and outline some key aspects that would be important to consider when approaching future projects within the smart manufacturing domain.

5.1 Preparation for future smart manufacturing

According to [Schuh et al., 2017], the process of migrating to I4.0 is defined by moving further up the maturity index, as explained in section 3.1.3. To summarize, the areas in the model are *computerization*, *connectivity*, *visibility*, *transparency*, *predictive capacity* and *adaptability*. While the process of moving up the I4.0 maturity index looks linear, it would be better to think of it from an iterative viewpoint. Every step in the model is built upon its predecessor, and a more sophisticated foundation enables more advanced insights and analytics.

Because of this, the *computerization* and *connectivity* steps should be revisited with every improvement of development in the smart factory foundation, in order to further optimize and enable new functionality, see figure 5.1.

As mentioned, the analysis in this thesis will primarily focus on ideas regarding the basic structure of IT and connectivity. It will, however, also mention more sophisticated, higher level analytics, but not in particular depth.

Within this area, an abundance of interesting and highly relevant analyses could be conducted. The thesis has categorized the relevant focus areas into *company wide development architecture*, *information modeling*, *communication structures*, *computational modules* and *collaboration with other companies and organizations*. These areas will be analyzed in sections 5.2 to 5.6.

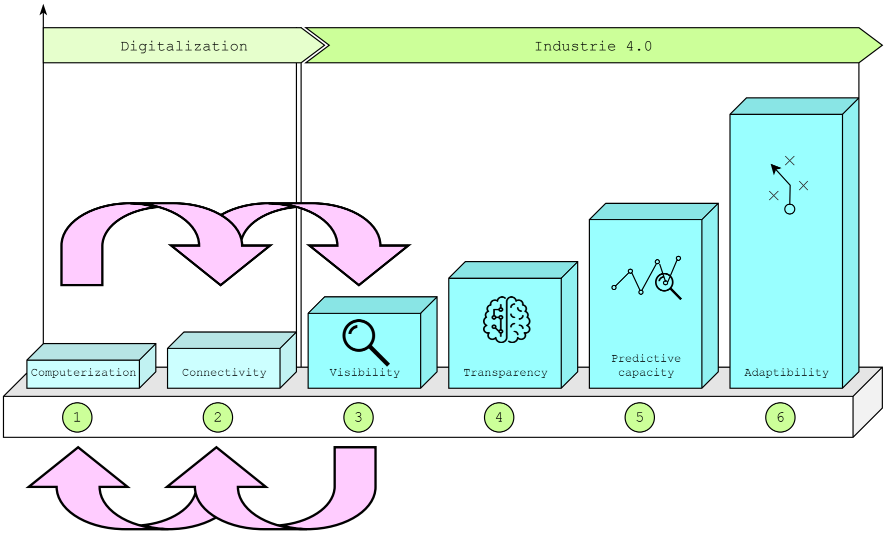

Figure 5.1 Revised representation of Industrie 4.0 maturity index

5.2 Company wide development architecture

One of the most important aspects of introducing changes and improvements into a manufacturing workflow is to clearly outline *what* the changes are trying to accomplish and *why* this is should be a priority for the parties concerned. The process of moving into a more data driven workflow creates a strict demand for cross-disciplinary competences, specifically where experts from a manufacturing background need to collaborate with IT-competent colleagues as well as data scientists, in companies where these disciplines are somewhat separated.

In order for this coalition of different competences to be successful, IT-architects and data scientist need to acquire a sufficient background and presence in the manufacturing domain. The experts at the production plant also need to adopt a sense of urge to what problems and improvements that could be addressed using more advanced IT-structures and analytics within the field.

Most modern manufacturing companies have adopted some sort of ML strategies into their future goals. The advancements are, however, still at a relatively small scale, and far from systematically incorporated into the actual workflow of most manufacturing facilities. It is undoubtedly important to conduct small to medium scale experiments and projects within the production lines, but a company wide development strategy must be formulated, especially with regard to the information and knowledge infrastructure surrounding data driven analytics. For a successful transition into this future workflow, the data driven vision must be appropriately anchored across the entire company, with heavy emphasis on the operational production crew.

Ultimately, this vision can be applied to each individual project and process, and the knowledge and experience gained should be easily accessible to anyone interested in further researching the subject. Of course, the difficulty for incorporating a transition like this increases with the size of the company, but can be made possible with Scania's resources and industry leading manufacturing experience.

5.2.1 Method for development communication

When working with the development of I4.0 with such a large company as Scania, it is important to put the discussions in the correct context. This is where an architecture as RAMI 4.0 could prove helpful. RAMI 4.0 aims to identify *what* aspect, within the ever more complex domain of I4.0, of the company that is being discussed. Such an architecture could aid the architects and engineers in identifying overlapping systems or functionalities which otherwise could be difficult to find. If every aspect is put in the context of where it fits within the RAMI 4.0, the company as a whole could potentially get rid of waste in the production process, by finding redundant or even competing functionalities. This could, for example be two different analytics systems aiming to solve the same problem, which could create confusion and ultimately hinder development. The architecture is designed to work on a company wide lever, covering all aspects of the process.

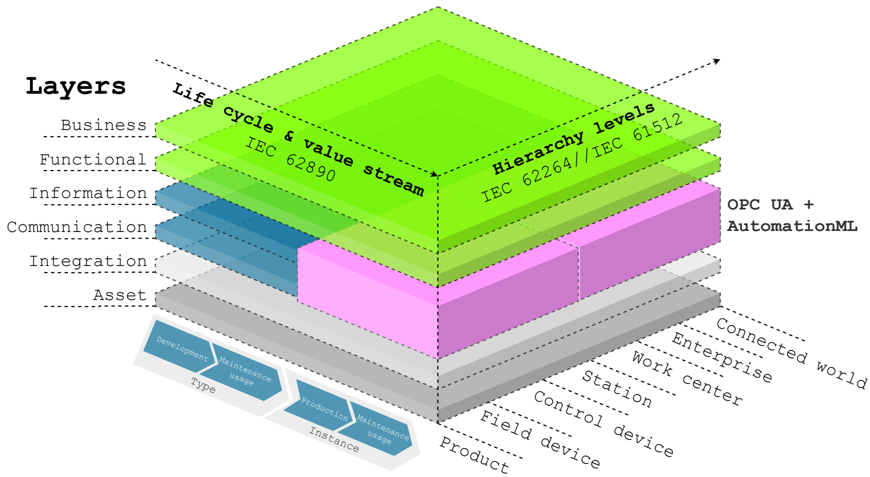

Figure 5.2 Open Platform Communications Unified Architecture in Reference Architecture Model Industrie 4.0

The use of a uniform reference architecture for development and mapping of functionality is especially important when working with external organizations and institutions, where terminology and best practices may differ.

5.2.2 Project suggestion for working with RAMI 4.0

For clarification, one widely adopted communication protocol within Scania, namely OPC UA has been mapped within RAMI 4.0. Here the entire communication and information layer of RAMI 4.0 can be implemented with OPC UA, together with AutomationML, as seen in figure 5.2.

There are some observations that could be given some extra attention. First, the OPC UA protocol, together with AutomationML, can be used across the entire company, i.e. from the product, through the entire manufacturing process[1] up to the, so called, connected world.

Another major observation is how the *instance-type* is used in this case. Of course, when OPC Foundation are working with OPC UA as a concept, the communication protocol is considered a *type*. However, when a company, such as Scania, implements the protocol, it instead will be considered an *instance*. All necessary documentation from the design process will still be available inside the AAS of OPC UA, thus facilitating the implementation and service process for the customer, which in this case is Scania.

The project should also make use of the CP classification, mentioned in section 3.1.7.1. This method of systematically and correctly classifying assets and the capabilities is crucial for an accurate representation inside RAMI 4.0 to be made.

The same analogy could be used by Scania across all different areas within the company. From this, a project could be initiated, with the objective of finding a framework for how to work with RAMI 4.0. This includes finding and defining certain routines for the employees to follow. Specific attention should be addressed to the life cycle and value stream of an asset within RAMI 4.0. This tool of systematic inclusion of documentation and operating information about an asset, as well as maintenance information to service technicians is valuable to both customers and manufacturers.

5.3 Information modeling

How information should be modeled in future factories is also an important step towards a smart manufacturing workflow. Because of rapid changes in structures and capabilities of information producing and consuming assets, the choice of modeling methods should be as generic as possible. Too complex and specified modeling approaches will obstruct the implementation process and could be at risk of quickly becoming obsolete. In regard to this, a simple and highly flexible modeling technique is required, that could be used across a vast majority of assets in the production plant.

[1] Low lever, realtime critical tasks could possibly be realized with the use of umati, which is built upon OPC UA

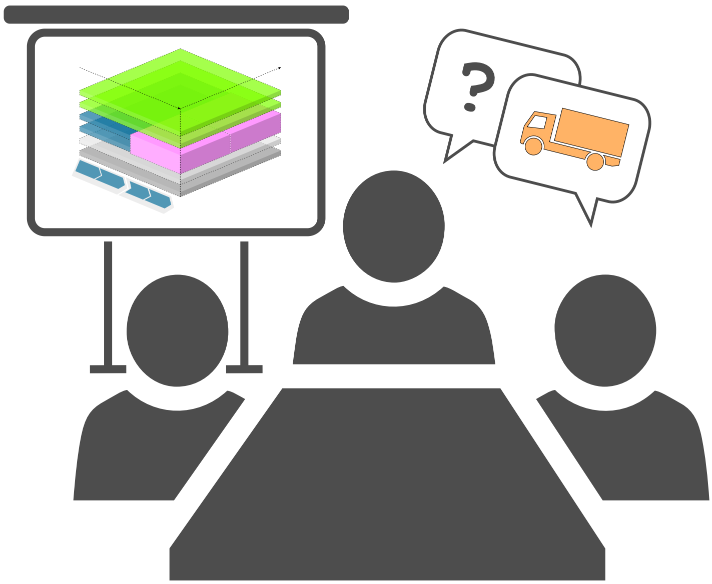

Figure 5.3 Project suggestion for Reference Architecture Model Industrie 4.0

5.3.1 Generic asset modeling

One example of such a modeling approach is that of the IoT platform ThingWorx from PTC, where each Thing, or more generally; asset, is modelled based on four different aspects, namely *properties*, *services*, *events* and *subscriptions* [PTC, 2019].

Properties and *services* could possibly be realized with AutomationML, within their AAS, here depicted in figure 5.4. The properties could describe their location within the production plant with CAEX, references to the ”sub-assets” belonging to the asset, geometry and kinematics with COLLADA, behavior and sequencing with PLCopen XML. Other formats could and should also be included, such as documentation, service, maintenance information, etc. The general structure of AutomationML can be seen in figure 5.5.

Events are created either internally or externally. It could be that an asset is beginning machining of a specific job, or that the machining process is finished. It could also be that some sort of anomaly or external disturbance is detected and the machining process is stopped.

Lastly, *subscriptions* are what connects events to services, either within the asset itself (e.g. a data change event triggered by a sensor could cause the subscribing asset to change its properties) or between assets (e.g. a stop in the manufacturing chain could cause the smart factory to reallocate resources and reschedule jobs).

This way of thinking about modeling of assets proves helpful when considering a dynamic communication structure, as mentioned later, in section 5.4.

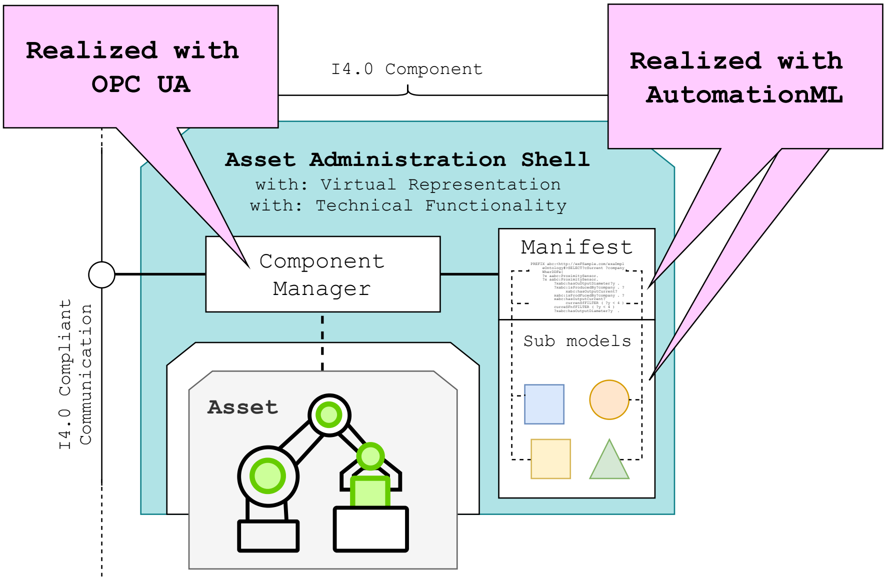

Figure 5.4 Realization of Asset Administration Shell with Open Platform Communications Unified Architecture and Automation Markup Language

5.3.2 Models for simulation

Generally, industrial machining assets are expensive and need significant consideration before investment. Simulation of the assets in a conceptual manufacturing line could greatly decrease the financial risk associated with such an investment. Defining models for simulation are, however, a time-consuming and possibly expensive task.

If the physical and logical models and processes for each asset is known beforehand, the process of simulation will become much more efficient. These models would be stored within the AAS and they can be created using COLLADA or PLCopen XML for instance. It is also worth mentioning the upcoming standard ISO 16400 for Equipment Behavior Catalogues (EBC), which specifically addresses the simulation potential through a standardized set of behaviors. At the time of writing this thesis, the standard has not yet been approved, but Scania should closely follow its development process, to determine whether the suggested standard could prove useful or not.

5.3.3 Data models

Another subject that also should be considered, with regard to the massive amounts of data that is being produced in modern factories, is how this data should be formatted. It is not uncommon that different data producing assets from different suppliers use different structures and format on their data.

To put this in context, a red light status for a certain machining asset, does not necessarily indicate the same condition as a red light on some other machining asset. In this example, the inconsistency in reporting statuses aggravates the process of designing generic analytics modules. One particular area of interest would be the standardized methods for I4.0 semantics, which has the ambition of creating a ”common language” for industry applications.

If the setup process could be facilitated, and the translation threshold for different assets could be eliminated, more resources could be made available to develop sophisticated analytics methods.

5.3.4 Location of AAS

One question that may come up when discussing AAS is where this digital representation of the physical object actually is stored. The quick answer is that it does not really matter. The importance lies instead in how it is accessed by other assets.

If the asset has some sort of internal repository capabilities, while also having an I4.0 compliant communication ability, then the AAS could be stored locally on the asset. The shell could, however, also be stored in some central (or semi-central) repository, which is also connected to the I4.0 network.

In order for other assets to access the state, properties and services of the asset in question, through its AAS, it only needs some globally unique identification. In practice, this could be an internal naming system, proceeded by a unique domain name connected to the internet.

5.3.5 Project suggestion for generic AAS

Concluding this, one potentially interesting development would be to implement AAS on a system-wide level. This could be done with a project or thesis work, with the goal of creating a uniform AAS (with modifiable properties) for hardware components (e.g. machining assets or robots), pieces of computational software (e.g. ML modules or controllers), human assets (e.g. operators or team managers) and products (e.g. engines or trucks), see figure 5.6. The CP classification mentioned in section 3.1.7.1 should also be taken into consideration for these assets. The goal with such a project would be to turn the relatively abstract idea of administrating all assets in a similar fashion into an actual, implementable solution.

For this project to be successful, one need to study AutomationML in more detail and determine what aspects could prove helpful when modeling properties and functionality.

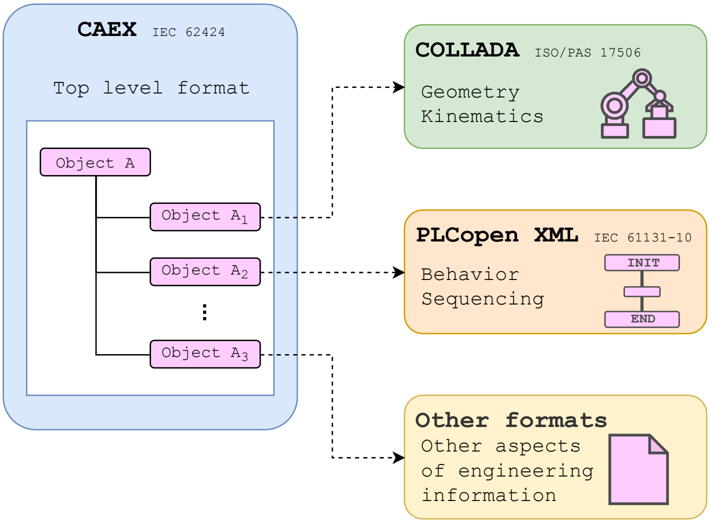

Figure 5.5 Automation Markup Language with different modeling components

5.4 Communication structures

One key aspect of the transition to I4.0 is how the communication across the company as a whole should be structured. As mentioned in section 3.2.4, the recognized structure of current manufacturing plants is strictly hierarchical, where each level communicates exclusively with the levels above and beneath. The field devices communicate with the PLCs, which in turn communicate with the process control level and so on.

5.4.1 Centralization and decentralization

One common solution inside a manufacturing plant is to direct most of the communication through some bus communication, like Production Service Bus. In future scenarios, it would be interesting to extend the more traditional centralized communication structure to a more decentralized approach, as seen in figure 3.9. This could greatly decrease the needed bandwidth and allow for improved scalability in the plant. Additionally this could reduce the response time between edge devices, which in turn could enable more efficient control applications.

For this to be useful, it is necessary to develop new, dynamic communication structures with simple installation and configuration. Ideally, a plug-and-play configuration would be preferable. This kind of functionality further drives the need for a standardized approach for systems from a variety of Original Equipment Manufacturers (OEMs) and machine tool suppliers to seamlessly connect and operate.

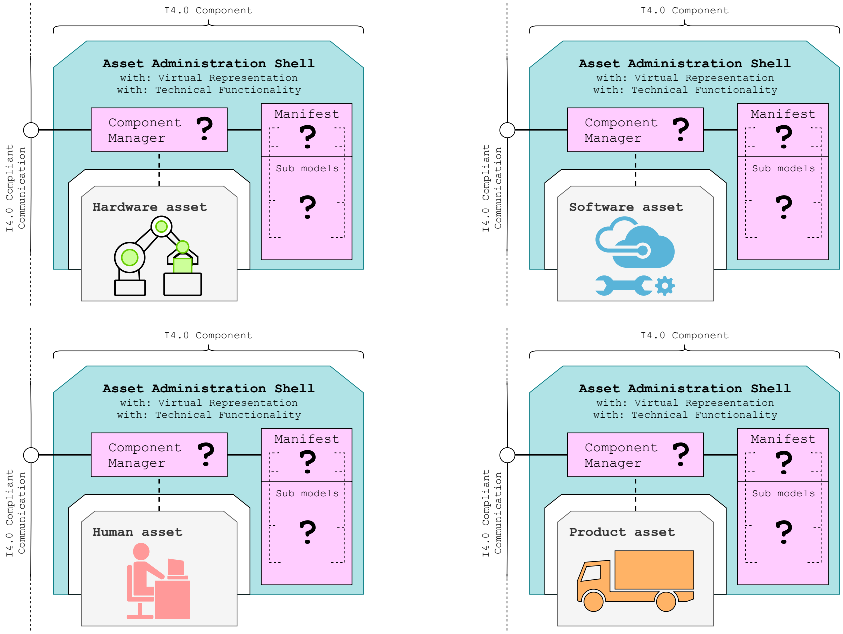

Figure 5.6 Project suggestion for generic Asset Administration Shell

This will, however, not be an easy task to accomplish, and any party that seeks to have influence over the development actively has to participate in the ongoing global discussions in the matter. From a more local point of view, it is important for manufacturing companies to explore different operation methods and build a substantial knowledge base regarding what to prioritize during these discussions. This could include what communicational protocols that should be used, what type of technology the network should be based on, etc.

5.4.2 Dynamic initiation of assets

Dynamic and simple initiation of a new asset into some smart, I4.0 compliant communication protocol needs some basic fundaments to work. First, it must be made simple for the asset to connect to the network. This would preferably be through some wireless network, like the upcoming technology 5G. Additionally, the asset needs to cary information about itself and how it operates. This would ideally be included in its AAS. This information should include, but not be exclusive to:

- a globally unique identifier
- information regarding what communication protocols the asset can use
- detailed information regarding the asset's operational capacity, function, etc

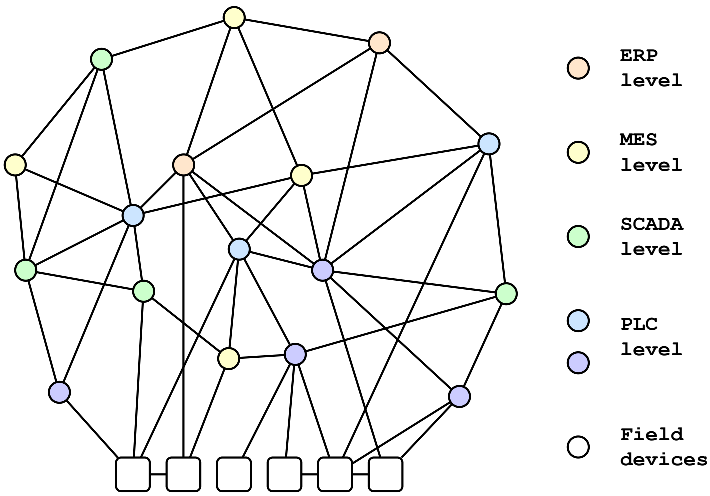

Figure 5.7 Representation of communication cluster

- details regarding what data the asset is able to produce
- its active connections to other assets

When the asset is physically connected to the network, the preferred scenario would be if the asset itself could provide the network the information on its communication capabilities, such as high and low level transfer protocols. This will enable the network to correctly initiate the communication with the asset. When the connection is established, the asset will provide a description of its functionality as well as what kind of data the asset will produce. This description could then help the other assets in the network to decide, manually or automatically, whether or not to subscribe to certain events. The interconnection between the communicative and descriptive capabilities of the assets is of great importance, and it would be favorable to choose these frameworks in conjunction with each other. One such combination could be OPC UA and AutomationML, which are engineered to work together as seamlessly as possible.

5.4.3 Project suggestion for dynamic communication structure

Following the uncertainty regarding how to dynamically and automatically initiate new devices and nodes into the network, a project regarding this matter could lead to great advancements for smart manufacturing. The project could focus on how the asset, using the information and properties provided in its AAS, could state its participation in the network. Other assets could then analyze the newly connected asset's properties to decide wether or not the new node is of interest and in that case initiate a more direct communication. Following the reasoning in section 5.3.1, the interested assets could then automatically create a subscription to certain events from the connected assets and even initiate services depending on their task. This has been illustrated in figure 5.8.

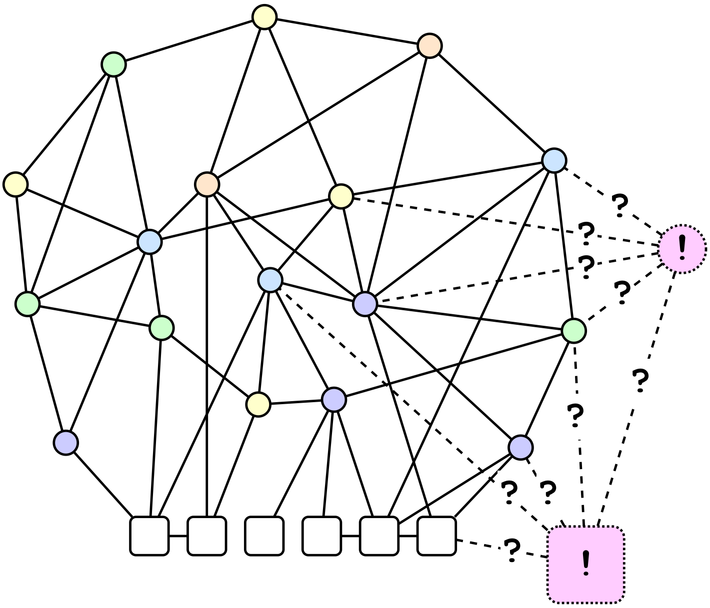

Figure 5.8 Project suggestion for dynamic communication structures

5.5 Computational modules

Following the same analogy as in section 5.2.1 and 5.3.1, any computational software modules should also be considered as assets, and also be surrounded with an AAS. These modules should also, preferably, be introduced into the network as described in section 5.4. The modules then need to be crafted depending on what problem they are addressing. A module for predictive maintenance with data from sequential physical processes could maybe use RNNs or LSTMs, mentioned in section 3.4.3.3. Another module for automatic quality control and assurance using cameras and optical sensors could maybe instead use CNNs to solve the task.

There are also many different unsupervised problems that could be addressed, which conceptually are more difficult to incorporate. One suggestion including such an unsupervised algorithm will follow in section 5.5.3.

5.5.1 Non-disruptive online modules

A common case in modern manufacturing sites is that data are extracted from the concerned systems, processed and analyzed and eventually presented as decision support. This is an example of an offline module, which may also interrupt the systems from where the data are gathered. The perspective on data gathering should be that offline data is considered outdated the moment it is extracted from the system.

A more ideal scenario would be for the computational analytics module to subscribe to events, thus not interrupting nor alerting the observed systems in any way. This preferred scenario would also include realtime, online updates to the analytics engine.

Of course, to implement realtime applications from zero position could be an overwhelming task, so smaller experimental projects become a necessity. These projects should, however, always be engineered with online functionality in mind.

5.5.2 Feedback from modules to physical world

The question about how these non-disruptive systems should feedback information or actions back into the network should also be addressed. Ideally, the data producing assets are also given the opportunity of subscribing to the events created by the computational modules. These subscriptions then trigger some service which alters the operations routing of the asset.

5.5.3 Project suggestion for computational modules

It would be exiting to initiate a project based on the research made by [Bagheri et al., 2015], as mentioned in section 3.4.4. As explained, the computations within the method is supposed to take place in the *cyber layer* from section 3.2.3. The project would then include letting the module build a knowledge base of different states, based on the operating routines of the assets. The assets could compare their performance to ”normal” states and then possibly predict the time to next failure.

Such a project will need some advanced unsupervised clustering algorithms to be successful. Examples of techniques that could be used are dimensionality reduction using PCA and k-means clustering. There are also some more advanced techniques involving neural networks. The initial part of the project would, however put heavy emphasis on outlining the actual objective using the data at hand, before attempting to choose a correct ML method.

In these initial stages of AA, much more focus should be put into enabling the infrastructure for computational modules to be used, rather than working with optimizing performance for specific algorithms. This could, for a future project in the area, mean that the computational module itself could be imported from some other company, and that the objective would be to implement said module in an efficient manner.

5.6 Collaboration with other companies and organizations

This section will include some general observations and conclusions without connection to any specific theoretical model.

As mentioned in previous sections, this is an area in which there lies huge potential for the future of manufacturing. Because of the complexity of modern manufacturing facilities and increasing need for external systems and services, the industrial companies of the future need to adopt their perspective on competition and collaboration. It will be of increasing importance to make the production and development processes more transparent, both internally to other divisions within the companies and to external stakeholders, such as suppliers, customers, organizations, etc. This rising demand is, however, not unidirectional. Scania, as a company of significant influence, should also require transparency from its stakeholders and thus leading the innovative change.

This is concretely done by requiring thorough and updated documentation and openness from ordered machining tools and systems, as well as providing it for the following parties in the value chain. All of this fits within the boundaries of the RAMI 4.0 and AAS, where every aspect of an asset's lifecycle can be managed in a uniform manner. In order for this to be made possible, the need for joint working groups between companies and global discussions with standardization organizations and organizations arises.

The future of industrial manufacturing is already being crafted, and ultimately it comes down to whether to lead the innovation development or to be a passenger in the process. Scania is, in many regards, industry leading with their state-of-the-art approach to modularity amongst their product line, where an almost countless number of products can be produced from a highly limited array of articles. This idea should be extended into the manufacturing domain, in order to stay competitive in the future of smart manufacturing.

5.6.1 PTC

One company that potentially could help accelerate the development of smart manufacturing would be PTC. PTC provides software and solutions for IoT modeling with ThingWorx, AR modeling with Vuforia Studio and data analytics with ColdLight. These pieces of software could enable generic modeling, as discussed in section 5.3.1. One next step could be to investigate how PTC's software could be combined with the European standards described in the thesis and potentially integrate PTC's services into Scania's operations.

6

Discussion

This chapter will evaluate and discuss the findings made in the thesis, with respect to the research questions.

6.1 Evaluation of results

The analysis chapter of this thesis work has brought up some ideas of ways to proceed in order to drive AA into manufacturing. The project ideas presented offer different degrees of abstraction, and this thesis work offers little suggestion on how to actually implement this is real world applications, but offers instead a wider picture of goals that are more long term. In order to dive down into details and solutions for implementation, a more specific case formulation is required. The thesis should, however, not be regarded as a step-by-step implementation manual, but rather as a study of the ongoing developments in the smart manufacturing scene.

Next, the research questions for the thesis will be reviewed. A short discussion regarding how well the thesis managed to answer the questions will be presented, as well as some notes for the aspects in where the thesis may have failed.

6.1.1 First research question

The first research question adressed a more general view of smart manufacturing and was formulated as follows.

- What actions should be performed to drive the integration of Advanced Analytics into a manufacturing workflow?

The thesis successfully manages to provide answers regarding where to focus future research at Scania. On a more abstract level, the single most important objective to pursue is to prepare the company for a future involving closer collaboration with standardization organizations and collaboration partners, as well as with competitors. In order for this to be made possible, a company-wide architecture for identifying functionality and information structures needs to be introduced. There are some different approaches world wide, but the one focused on in this project has been RAMI 4.0.

Other notable actions would be to introduce some sort of generalized way of uniformly modeling assets across the entire company. Again, this is preferably done using standardized models, such as AAS together with OPC UA and AutomationML, as mentioned in previous chapters. There are different existing solutions for this kind of problems, many of which this thesis fails to mention. Those solutions mentioned are, however, sensibly as generic as possible.

Furthermore, more research needs to be conducted regarding the changing requirements of modern communication structures. The conclusions made in the thesis regarding these structures are rather vague, but hopefully offer new perspectives and ideas for smart manufacturing.

Many of the standards suggested in the thesis are currently under development, and may not be ready for full implementation. The developments are, however, in rapid growth, and it will be crucial to follow these standards as they potentially will shape the future of modern manufacturing facilities on a global scale.

6.1.2 Second research question

The second research question was more specifically addressed towards analytics and machine learning, and was formulated as follows.

- What would be the ideal Advanced Analytics methodologies for usage in different manufacturing applications?

The second question turned out to be more difficult to address in a generalized way. There are many different ML algorithms, and they need to be assessed differently for each individual use case.

One potential problem associated with using ML inside manufacturing systems is that the ”black box” approach may not always be preferred. Many times, engineers and system designers need control and insight into the different algorithms within their manufacturing processes, which may be contradictory to the ideas presented with ML.

Another problem is that many of the areas that hopefully could be addressed using ML techniques, such as predictive maintenance, simply have too little data for the models to obtain sufficient training. Even if the assets produce continuous manufacturing data, they might not fail often enough for the algorithms to recognize patterns. This could, of course, be facilitated by closer collaboration with manufacturers and customers, where more data could be made available.

In summary, this research question has not been answered in as much depth as initially anticipated. This in mostly due to that developments regarding the first research question states the requirements for answering the second.

6.2 Credibility of the study

One aspect that might affect the study's credibility negatively is the fact that the author of the thesis had somewhat limited knowledge beforehand about manufacturing processes in practice. One way of addressing this potential lack of knowledge has been to put the recommended lines of action on a more conceptual level.

Furthermore, a considerable amount of material used in this study originates from standardization organizations and projects, where different large companies are in cooperation with each other. Many of these standards are open, and their ambitions is to truly create a uniform future manufacturing scene. The aspect of openness and standardization has carefully been considered throughout the process of the thesis work, which enhances the credibility of the study.

The standards discussed in the thesis include:

- **IEC 62264** for Enterprise-Control System Integration in ISA-95
- **IEC 62890** for Life cycle & value streams
- **IEC 62541** for Open Platform Communications Unified Architecture
- **IEC 62714** for Automation Markup Language
- **IEC 62424** for Computer Aided Engineering Exchange
- **ISO/PAS 17506** for COLLAborative Design Activity
- **IEC 61131-10** for PLCopen XML
- **ISO 16400** for Equipment Behavior Catalogues

7 Summary

This chapter will summarize the thesis, as well as present the recommendations for future research.

To conclude this thesis work, a number of interesting areas within Industrie 4.0 (I4.0) have been investigated. These areas include the use of Cyber-Physical Systems (CPSs) in manufacturing facilities, the enormous increase in data volumes generated from modern assets and Machine Learning (ML) methodologies and algorithms for different applications.

Most focus has been put into the infrastructure for future smart manufacturing facilities, and the advancements that this development could convey. After the literature review had been conducted within the areas mentioned above, the analysis was made, divided in five different focus areas. Following this, four of the areas included suggestions for future work, as summarized in 7.1.

The first of the five areas outlined a framework for incorporating a company wide development architecture, as well as some key factors to consider when working with it. This included the use of cross-discipline collaboration within working groups with clearly defined visionary goals to create a sense of urgency for involved parties. Lastly, this section treated the incorporation of Reference Architecture Model Industrie 4.0 (RAMI 4.0) for development processes, which is a valuable tool for building I4.0 systems.

Next, some concepts regarding information processes were discussed, along with generalized methods for treating assets and their functionalities. These concepts included an array of standardized modells and technologies such as Asset Administration Shell (AAS), Open Platform Communications Unified Architecture (OPC UA), Automation Markup Language (AutomationML), Computer Aided Engineering Exchange (CAEX) and COLLAborative Design Activity (COLLADA). Additionally, section 5.3 shortly treated simulation of processes with Equipment Behavior Catalogues (EBC) and basic ideas regarding uniform I4.0 semantics.

Proceeding this, some ideas were brought up regarding communication structures of future manufacturing facilities. This included the decentralizing of higher level production systems, in order to decrease the hierarchical structure found in many modern companies. This (semi-)decentralization could enable for more efficient communication of time-sensitive tasks. As a requirement for this, the thesis states that new methods for dynamic initiation need to be researched and further explored.

The thesis also discusses some aspects surrounding different computational modules for analytics. These modules should also be considered assets within the same framework as mentioned above, and they should be of non-disruptive nature. The modules should always be developed with online capabilities in mind. Additionally, a need for standardized interfaces arises when importing solutions from external parties, in order for the modules to be integrated in a seamless fashion.

The last section in the analysis outlined some thoughts about the future of collaboration between companies and organizations, which will accelerate the evolution of smart, connected manufacturing facilities. The thesis suggests that this will be accomplished by increased transparency of the production processes.

7.1 Future work

As this thesis was of investigative nature, further, more elaborative research on the ideas covered needs to be conducted. From the analysis in chapter 5, four distinct project ideas have been suggested, namely

Project suggestion 1: Framework for using RAMI 4.0

Project suggestion 2: Generic modeling of AAS

Project suggestion 3: Dynamic communication structures

Project suggestion 4: Computational modules

These projects could be initiated in parallell, however, Scania should preferably prioritize the routines surrounding RAMI 4.0 before attempting the latter projects. Another note that should be stated is that, however tempting it appears to work with computational modules for analytics, they require the information infrastructure from asset modeling and dynamic communication structure to truly be successful.

Additionally, while not mentioned as a project suggestion, it is of great importance to initiate collaborative efforts with other companies and organizations, as the developments shaping the future in smart manufacturing will occur on a global scale.

Bibliography

Adolphs, P, S Auer, M Billmann, M Hankel, R Heidel, M Hoffmeister, H Huhle, M Jochem, M Kiele, G Koschnick, et al. (2016). *Structure of the Administration Shell*. [Online; accessed 20-September-2018]. Federal Ministry for Economic Affairs and Energy (BMWi), Berlin. URL: https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/structure-of-the-administration-shell.pdf?__blob=publicationFile&v=7.

Arm, J., F. Zezulka, Z. Bradac, P. Marcon, V. Kaczmarczyk, T. Benesl, and T. Schroeder (2018). "Implementing Industry 4.0 in Discrete Manufacturing: Options and Drawbacks". *IFAC-PapersOnLine* **51**:6. 15th IFAC Conference on Programmable Devices and Embedded Systems PDeS 2018, pp. 473 –478. ISSN: 2405-8963.

Bagheri, B., S. Yang, H.-A. Kao, and J. Lee (2015). "Cyber-physical Systems Architecture for Self-Aware Machines in Industry 4.0 Environment". *IFAC-PapersOnLine* **48**:3. 15th IFAC Symposium on Information Control Problems in Manufacturing, pp. 1622 –1627. ISSN: 2405-8963.

Baur, C. and D. Wee (2015). "Manufacturing's next act". *McKinsey Quarterly, Jun.* [Online; accessed 22-August-2018]. URL: https://www.timereaction.com/papers/manufacturing_next_act.pdf.

Bishop, C. (2006). *Pattern recognition and machine learning*. Vol. 4. Springer New York. URL: http://scholar.google.com/scholar.bib?q=info:jYxggZ6Ag1YJ:scholar.google.com/&output=citation&hl=en&as_sdt=0,5&as_vis=1&ct=citation&cd=0.

Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep Learning*. [Online; accessed 7-June-2019]. MIT Press. URL: http://www.deeplearningbook.org.

Herterich, M. M., F. Uebernickel, and W. Brenner (2015). "The Impact of Cyber-Physical Systems on Industrial Services in Manufacturing". *Procedia CIRP* **30**. 7th Industrial Product-Service Systems Conference - PSS, industry transformation for sustainability and business, pp. 323 –328. ISSN: 2212-8271.

Hoffmeister, M. (2015). *Industrie 4.0: The Industrie 4.0 Component*. ZVEI - German Electrical and Electronic Manufacturers' Association. URL: https://www.zvei.org/fileadmin/user_upload/Themen/Industrie_4.0/Das_Referenzarchitekturmodell_RAMI_4.0_und_die_Industrie_4.0-Komponente/pdf/ZVEI-Industrie-40-Component-English.pdf.

Iqbal, R., F. Doctor, B. More, S. Mahmud, and U. Yousuf (2017). "Big Data analytics and Computational Intel ligence for Cyber–Physical Systems: Recent trends and state of the art applications". *Future Generation Computer Systems*. ISSN: 0167-739X.

Johnsson, C. (2003). *ISA 95 - how and where can it be applied?* ISA - The Instrumentation, Systems and Automation Society.

Lee, J., B. Bagheri, and H.-A. Kao (2015). "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems". *Manufacturing Letters* **3**, pp. 18 –23. ISSN: 2213-8463.

Monostori, L. (2014). "Cyber-physical Production Systems: Roots, Expectations and Research and Development Challenges". *Procedia CIRP* **17**. Variety Management in Manufacturing, pp. 9 –13. ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2014.03.115. URL: http://www.sciencedirect.com/science/article/pii/S2212827114003497.

Monostori, L., B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, and K. Ueda (2016). "Cyber-physical systems in manufacturing". *CIRP Annals* **65**:2, pp. 621 –641. ISSN: 0007-8506.

Negri, E., L. Fumagalli, and M. Macchi (2017). "A Review of the Roles of Digital Twin in CPS-based Production Systems". *Procedia Manufacturing* **11**. 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy, pp. 939 –948. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2017.07.198. URL: http://www.sciencedirect.com/science/article/pii/S2351978917304067.

OPC Foundation (2017). "AutomationML, OPC UA, and the Asset Administration Shell of Industrie 4.0 Components". [Online; accessed 22-May-2019]. URL: https://opcfoundation.org/wp-content/uploads/2017/11/OPCUA-AutomationML-2017-v3.pdf.

Palmer, M. (2006). *Data is the new oil*. [Online; accessed 23-May-2019]. URL: https://ana.blogs.com/maestros/2006/11/data_is_the_new.html.

Paolanti, M., L. Romeo, A. Felicetti, A. Mancini, E. Frontoni, and J. Loncarski (2018). "Machine learning approach for predictive maintenance in industry 4.0". In: *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pp. 1–6. DOI: 10.1109/MESA.2018.8449150.

PTC (2019). *PTC: Enabling Industrial Digital Transformation*. [Online; accessed 7-June-2019]. URL: https://www.ptc.com/en/about/.

PTC (2019). *ThingWorx Thing Model Visual Representation*. [Online; accessed 7-June-2019]. URL: http://support.ptc.com/cs/help/thingworx_hc/thingworx_6.6_hc/index.jspx?id=ThingWorxThingModelVisualRepresentation&action=show.

Qi, Q. and F. Tao (2018). "Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison." *IEEE Access, Access, IEEE*, pp. 3585–3593. ISSN: 2169-3536.

Rüßmann, M., M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, and M. Harn isch (2015). "Industry 4.0: the future of productivity and growth in manufacturing industries". *Boston Consulting Group* **9**.

Scania (2019a). *Our solutions*. [Online; accessed 26-May-2019]. URL: https://www.scania.com/group/en/our-solutions/.

Scania (2019b). *Scania at a glance*. [Online; accessed 26-May-2019]. URL: https://www.scania.com/group/en/scania-at-a-glance/.

Schuh, G., R. Anderl, J. Gausemeier, M. ten Hompel, and W. Wahlster, (Eds.) (2017). *Industrie 4.0 Maturity Index: Managing the Digital Transformation of Companies*. acatech Studie. Utz, München. ISBN: 978-3-8316-4613-5. URL: http://www.acatech.de/de/publikationen/empfehlungen/acatech/detail/artikel/industrie-40-maturity-index-die-digitale-transformation-von-unternehmen-gestalten.html.

VDI/VDE (2016). *Status Report: Industrie 4.0 – Technical Assets Basic terminology concepts, life cycles and administration models*. VDI/VDE GMA Technical Committee FA7.21.

VDI/VDE (2017). *Industrie 4.0 Begriffe/Terms, VDI - Statusreport*. VDI/VDE GMA Technical Committee FA7.21.

Wagner, C., J. Grothoff, U. Epple, R. Drath, S. Malakuti, S. Grüner, M. Hoffmeister, and P. Zimermann (2017). "The role of the industry 4.0 asset administration shell and the digital twin during the life cycle of a plant". In: *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8. DOI: 10.1109/ETFA.2017.8247583.

Yan, J., Y. Meng, L. Lu, and L. Li (2017). "Industrial big data in an industry 4.0 environment: challenges, schemes, and applications for predictive maintenance". *IEEE Access* **5**, pp. 23484–23491. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2765544.

Zezulka, F., P. Marcon, Z. Bradac, J. Arm, T. Benesl, and I. Vesely (2018). "Communication Systems for Industry 4.0 and the IIoT". *IFAC-PapersOnLine* **51**:6. 15th IFAC Conference on Programmable Devices and Embedded Systems PDeS 2018, pp. 150 –155. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2018.07.145. URL: http://www.sciencedirect.com/science/article/pii/S2405896318308899.

Lund University
Department of Automatic Control
Box 118
SE-221 00 Lund Sweden

Document name
MASTER'S THESIS

Date of issue
July 2019

Document Number
TFRT-6089

Author(s)
Axel Lundholm

Supervisor
David Chauca, Scania
Charlotta Johnsson, Dept. of Automatic Control, Lund University, Sweden
Anders Robertsson, Dept. of Automatic Control, Lund University, Sweden (examiner)

Title and subtitle
Accelerate Business Value in Manufacturing with Advanced Analytics

Abstract
Recent developments in smart manufacturing enable convergence between the digital and physical worlds of modern manufacturing facilities. This evolution is, however, far from trivial and thorough research and investigation needs to be conducted regarding dynamic connectivity of assets and implementation of data driven analytics, which provides deeper insight into the operational processes.

Scania in Södertälje is the object for the case study, with the aim of presenting recommendations for future research projects within smart manufacturing. Also, PTC in Boston, Massachusetts, has contributed with expertise and knowledge in the matter. Addressing the problems regarding what future actions to pursue and what methodologies to invest research in, this thesis base its analysis and discussion on recent research papers and documentation from international standardization organizations.

The analysis identified and categorized the present problems into company wide development architecture, information modeling, communication structures, computational modules and collaboration with other companies and organizations.

Ultimately, four different project recommendations are presented. The first suggestion includes development of a framework for using Reference Architecture Model Industrie 4.0 (RAMI 4.0) in company wide development and research, as a way of categorizing systems and functions. Secondly, a suggestion for generic modeling of assets was presented. Assets could be anything from machining tools, to analytics software and even operational personell. Thirdly, the thesis recommends that Scania investigates dynamic communication structures, which breaks the traditional hierarchical view on information infrastructure across the company. Lastly, a project regarding non-intrusive, online computational modules was discussed. This suggestion was, however, not in particular detail, as the thesis concluded that the foundation for data driven methods is of the highest importance, rather than suggesting actual analytics algorithms.

Keywords
Smart manufacturing, Industrie 4.0, Reference Architecture Model Industrie 4.0, Asset Administration Shell, Cyber-Physical Systems, Big Data, Advanced Analytics

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280-5316

ISBN

Language
English

Number of pages
1-53

Recipient's notes

Security classification

http://www.control.lth.se/publications/