A System Approach to Implementation of Predictive Maintenance with Machine Learning

by

Chen Ye

B.E. Measurement-Control Technology and Instrumentation
Harbin University of Science and Technology, 2007

SUBMITTED TO THE SYSTEM DESIGN AND MANAGEMENT PROGRAM IN PARTIAL FUFILLMENT OF THE REQURIEMENTS FOR DEGREE IN

MASTERS OF SCIENCE IN ENGINEERING AND MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2018

© 2018 Chen Ye All rights reserved.

The author herby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now or known hereafter created.

Signature redacted

Signature of Author: ______________________________

System Design and Management Program
May 29th, 2018

Certified by: ______________ Signature redacted

Dr. Donna H Rhodes
Principal Research Scientist, Sociotechnical System Research Center
Thesis Supervisor

Accepted by: ______________ Signature redacted

Joan S. Rubin
Senior Lecturer of Engineering Systems
Executive Director, System Design & Management Program

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 20 2018
LIBRARIES
ARCHIVES

MIT Libraries

77 Massachusetts Avenue
Cambridge, MA 02139
http://libraries.mit.edu/ask

DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

The images contained in this document are of the best quality available.

This page intentionally left blank.

A System Approach to Implementation of Predictive Maintenance with Machine Learning

by

Chen Ye

Submitted to the System Design and Management Program on May 29th, 2018 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management

ABSTRACT

Digital technology is changing the industrial sector, yet how to make rational use of some technologies and create considerable value in a variety of industrial scenarios is an issue. Many digital industrial companies have stated that they have helped clients with their digital transformation, create much value, but the real effects have not been shown in public. Venture capitals firms have made huge investment in potential digital industrial startups. Numerous industrial IoT platforms are emerging in the market, but a number of them fade soon after. Many people have heard about industrial maintenance technology, but they have difficulty in differentiate concepts such as reactive maintenance, planned maintenance, proactive maintenance, and predictive maintenance. Many people know that big data and AI are essential in industrial sector, but they do not know how to process, analyze, and extract value from industrial data and how to use AI algorithms and tools to implement a research project.

This thesis analyzes the entire digital industrial ecosystem in various dimensions such as initiatives, technologies in related domains, stakeholders, markets, and strategies. This work also analyzes of the predictive maintenance solution in various dimensions such as background, importance, suitable scenarios, market, business model, and technology. The author plans an experiment for the predictive maintenance solution, including goal, data source and description, methods and steps, and flow and tools. Then author uses a baseline approach and an optimal approach to implement the experiment, including data preparation, selection and evaluation of both regression and classification models, and deep learning practice through neural network building and optimization. Finally, contributions and expectations, and limitations and future research are discussed. This work uses a system approach, including system architecting, system engineering, and project management, to complete the process of analysis, design, and implementation.

Thesis Supervisor: Dr. Donna H Rhodes
Title: Principal Research Scientist, Sociotechnical System Research Center

Acknowledgements

First of all, I would like to express my gratitude to my thesis supervisor, Dr. Donna H. Rhodes, who gently accepted to supervise my thesis. Her great support, suggestion, and encouragement have been essential to my research. She even spent her personal time at weekend on reviewing my work. Without her full support, my graduate work would not be possible. It is my honor to have this wonderful opportunity to work with her and I hope to have more opportunities later.

Also, I would like to thank my advisors at SDM, Joan S. Rubin and William F. Foley for their suggestion and help in my wonderful academic and curricular experience in the two years at MIT. And I would like to thank all the SDM program faculty for the professional education on system thinking, which helped me lay a solid foundation on system architecting, system engineering, and project management. And I would like to thank all the SDM administration stuff for their supports and help since I was a prospective student.

I would like to thank my classmates, I will never forget the time all we studied and played together. I cannot remember how many nights we stayed up very late for our team work. Your assistance and encouragement helped me to go through my student life in the United States. I am proud of being one of the SDM'16 cohort.

I want to thank my family for their understandings and backup throughout this process. My parents, my wife Juan, my daughter Spica, and my newborn Mia are the most precious asset in my life. Without your assistance, I can do nothing.

Table of Contents

CHAPTER 1: INTRODUCTION **10**

1.1 PROBLEM STATEMENT 10

1.2 RESEARCH OBJECTIVES 12

1.3 RESEARCH APPROACH 12

CHAPTER 2: ECOSYSTEM ANALYSIS **14**

2.1 BACKGROUND ANALYSIS 14

2.1.1 Industry 4.0 .... *14*

2.1.2 Industrial Internet *16*

2.1.3 Made in China 2025 .... *19*

2.1.4 The Industrial Value Chain Initiative (IVI) *20*

2.1.5 Alternatives Initiatives *21*

2.2 TECHNOLOGY ANALYSIS 21

2.2.1 IT/OT Convergence *23*

2.2.2 M2M and Machine Protocols *25*

2.2.3 Convergence of H2H, H2M, and M2M *27*

2.2.4 Internet of Things (IoT) *29*

2.2.5 Network and Standard *31*

2.2.5.1 Time-Sensitive Networking (TSN) 33

2.2.5.2 OPC UA 35

2.2.5.3 TSN + OPC UA 35

2.2.6 Cloud Computing and Platform *37*

2.2.7 Edge Computing and Device *40*

2.2.8 Big Data and Analytics *43*

2.2.9 Artificial Intelligence (AI) and Machine Learning *45*

2.2.10 Other Technologies *48*

2.3 STAKEHOLDER ANALYSIS 48

2.4 MARKET ANALYSIS 50

2.4.1 Market Trend *50*

2.4.2 Market Analysis by Vertical *51*

2.4.3 Market Analysis by Product/Service *52*
2.4.4 Market Analysis by Geography *54*
2.5 STRATEGY ANALYSIS 55
2.5.1 Product vs Service *55*
2.5.2 Platform vs Application (Horizontal vs Vertical) *56*
CHAPTER 3: SOLUTION ANALYSIS 59
3.1 BACKGROUND OF MAINTENANCE 59
3.2 INTRODUCTION OF PREDICTIVE MAINTENANCE 61
3.3 MARKET ANALYSIS 63
3.4 BUSINESS MODEL ANALYSIS 64
3.5 TECHNOLOGY REVIEW 64
CHAPTER 4: EXPERIMENT DESIGN 67
4.1 GOAL 67
4.2 DATA SOURCE AND DESCRIPTION 67
4.3 METHODS, STEPS, AND TOOLS 70
CHAPTER 5: EXPERIMENT IMPLEMENTATION 72
5.1. BASELINE APPROACH 72
5.1.1 Data Preparation *72*
5.1.1.1 Data Reading 72
5.1.1.2 Data Labeling 72
5.1.1.3 Data Visualization 73
5.1.1.4 Data Normalization 76
5.1.2 Regression Model Selection and Evaluation *76*
5.1.3 Binary Classification Model Selection and Evaluation *80*
5.2 OPTIMAL APPROACH 81
5.3 FUTURE RESEARCH 82
CHAPTER 6: SUMMARY AND CONCLUSION 83
6.1 SUMMARY 83
6.2 CONTRIBUTIONS AND EXPECTATION 84
6.3 LIMITATIONS AND FUTURE RESEARCH 85

REFERENCES **87**

List of Figures

FIGURE 1. INDUSTRIAL IOT PROBLEM TAXONOMY (SOURCE: [1]) 11
FIGURE 2. ROADMAP OF INDUSTRIAL REVOLUTIONS (SOURCE: [4]) 15
FIGURE 3. DESIGN PRINCIPLES IN INDUSTRY 4.0 16
FIGURE 4. FUNCTIONAL DOMAINS (SOURCE: [7]) 18
FIGURE 5. MAPPING BETWEEN A THREE-TIER ARCHITECTURE TO THE FUNCTIONAL DOMAINS (SOURCE: [7]) 18
FIGURE 6. GOALS OF MADE IN CHINA 2025 (SOURCE: [9]) 19
FIGURE 7. 10 KEY AREAS THAT PROMOTE BREAKTHROUGHS (SOURCE: [9]) 19
FIGURE 8. LOOSE STANDARD IN THE REFERENCE MODEL IN IVI (SOURCE: [10]) 20
FIGURE 9. TECHNOLOGIES THAT TRANSFORM INDUSTRIAL PRODUCTION (SOURCE: [12]) 22
FIGURE 10. DIGITAL COMPASS (SOURCE: [13]) 23
FIGURE 11. CONVERGENCE OF IT/OT SYSTEMS (SOURCE: [17]) 25
FIGURE 12. M2M PROTOCOLS IN OSI MODEL (SOURCE: [19]) 26
FIGURE 13. CONVERGENCE OF H2H, H2M, AND M2M 29
FIGURE 14. DIFFERENCE BETWEEN M2M AND IOT (SOURCE: [22]) 30
FIGURE 15. TIMESCALE OF INDUSTRIAL COMMUNICATION STANDARDS (SOURCE: [25]) 32
FIGURE 16. INDUSTRIAL COMMUNICATION PROTOCOLS AND STANDARDS (SOURCE: [26]) 33
FIGURE 17. LAYER FOR TSN (SOURCE: [28]) 34
FIGURE 18. TSN CHARACTERISTICS 34
FIGURE 19. NEW ARCHITECTURE THAT FLATTENS THE AUTOMATION PYRAMID (SOURCE: [32]) 36
FIGURE 20. TRANSITION OF ARCHITECTURES FOR THE AUTOMATION (SOURCE: [33]) 36
FIGURE 21. THE ARCHITECTURE OF VARIOUS SERVICES IN PREDIX (SOURCE: [36]) 38
FIGURE 22. ARCHITECTURE OF PREDIX PLATFORM (SOURCE: [37]) 38
FIGURE 23. ARCHITECTURE OF COMPUTING FROM EDGE TO CLOUD (SOURCE: [39]) 41
FIGURE 24. CHARACTERISTICS OF EDGE COMPUTING USE CASE (SOURCE: [40]) 42

FIGURE 25. COOPERATION BETWEEN EDGE AND CLOUD (SOURCE: [40])..43
FIGURE 26. FOUR V'S OF BIG DATA (SOURCE: [41]) 44
FIGURE 27. AI TECHNOLOGIES CLASSIFICATION (SOURCE: [42]) 45
FIGURE 28. INDUSTRIAL AI WITH REQUIREMENTS AND CHALLENGES (SOURCE: [43]) 46
FIGURE 29. WORD CLOUD OF INDUSTRIAL AI ALGORITHMS (SOURCE: [45]) 47
FIGURE 30. STAKEHOLDER CLASSIFICATION BASED ON TECHNOLOGY (SOURCE: [46]) 49
FIGURE 31. IIOT REVENUE TREND (SOURCE: [47]) 50
FIGURE 32. IIOT MARKET SIZE AND IMPACT TREND (SOURCE: [48]) 51
FIGURE 33. MAIN INDUSTRIES IN IIOT MARKET (SOURCE: [50]) 52
FIGURE 34. MAJOR COMPANIES FOCUSED SECTORS 53
FIGURE 35. STARTUP FOCUSED IIOT STACKS (SOURCE: [51]) 53
FIGURE 36. INDUSTRIAL IOT MARKET BY REGION, 2022 (SOURCE: [52])....................................54
FIGURE 37. MAJOR IIOT PLAYERS BY GEOGRAPHY...54
FIGURE 38. MAJOR IIOT PLATFORM PROVIDER BY VERTICALS 56
FIGURE 39. MAGIC QUADRANT FOR INDUSTRIAL IOT PLATFORMS (SOURCE: [54])......................57
FIGURE 40. MAGIC QUADRANT FOR INDUSTRIAL IOT PLATFORMS (SOURCE: [54])......................58
FIGURE 41. MAINTENANCE STRATEGY CONTINUUM (SOURCE: [57])...60
FIGURE 42. COMPARISON IN OPERATION AMONG MAINTENANCE STRATEGIES (SOURCE: [58]).....61
FIGURE 43. THE IMPORTANCE OF PDM (SOURCE: [59])...62
FIGURE 44. PDM MARKET DEVELOPMENT (SOURCE: [58]) 63
FIGURE 45. DRIVING FORCES OF PDM MARKET DEVELOPMENT 63
FIGURE 46. PDM COMPANY RANKING (SOURCE: [58])..64
FIGURE 47. TRL CRITERIA (SOURCE: [62]) 65
FIGURE 48. SIMPLIFIED DIAGRAM OF ENGINE SIMULATED (SOURCE: [71]) 68
FIGURE 49. MODULES AND CONNECTIONS LAYOUT (SOURCE: [68]) 68
FIGURE 50. ONE GROUP OF TRAINING DATA 74
FIGURE 51. TEN GROUP OF TRAINING DATA 75
FIGURE 52. TEN GROUPS OF TRAINING DATA OVER RUL 76
FIGURE 53. FEATURE IMPORTANCE ANALYSIS 78

FIGURE 54. FEATURE COEFFICIENTS WEIGHTS 79
FIGURE 55. COST TRADEOFFS IN PREDICTION AND DETECTION (SOURCE: [1]) 86

List of Tables

TABLE 1. PROTOCOLS FOR M2M (SOURCE: [20]) 27
TABLE 2. FORM AND FUNCTION OF H2H, H2M, AND M2M 28
TABLE 3. COMPARISON OF VARIOUS IOT NETWORK (SOURCE: [24]) 31
TABLE 4. PRIMARY COMPONENTS OF PREDIX 40
TABLE 5. CLASSIFICATION OF EDGE 41
TABLE 6. PREDICATION OF NUMBER OF DEVICE THAT WILL BE CONNECTED 43
TABLE 7. RELATIONS AMONG FLOWS AND DOMAINS 49
TABLE 8. PRODUCT VS SERVICE (SOURCE: [53]) 55
TABLE 9. COMPARISON IN CONCEPT AMONG MAINTENANCE STRATEGIES 60
TABLE 10. SUITABLE AND UNSUITABLE SCENARIOS FOR PDM (SOURCE: [60]) 62
TABLE 11. STATISTICAL DATA (SOURCE: [63]) 68
TABLE 12. SENSORS' PHYSICAL MEANING (SOURCE: [68]) 69
TABLE 13. NAME OF EACH COLUMN IN THE RAW DATA (SOURCE: [63]) 70
TABLE 14. STEPS IN THE BASELINE APPROACH 71
TABLE 15. LABELED TRAINING DATA 73
TABLE 16. COMPARISON OF REGRESSION MODEL BASED ON PERFORMANCE METRICS 77
TABLE 17. RELATIONSHIP OF PERFORMANCE METRICS 80
TABLE 18. COMPARISON OF BINARY CLASSIFICATION MODEL BASED ON PERFORMANCE METRICS 81
TABLE 19. ARCHITECTURE OF THE NEURAL NETWORK FOR REGRESSION 81
TABLE 20. COMPARISON OF REGRESSION PERFORMANCE BETWEEN BASELINE AND OPTIMAL APPROACH 82

Chapter 1: Introduction

The digital age is here. We have seen and heard about cutting edge digital technologies such as Artificial Intelligence (AI), Big Data and Analytics, cloud computing, edge computing, Internet of Things (IoT), etc. Digital technology has powered a revolution of social network and consumer markets in the past two decades. Currently, digital technology is transforming the industrial sector. Imagine that aircraft engines will remind technicians of their overhaul, and steam turbines will communicate with each other to produce electricity more efficiently. These scenarios that happened in science fictions and movies will come to pass in the future. Digital technology makes it happen.

1.1 Problem Statement

Digital industry is here, yet how to make rational use of certain technologies and create considerable value in various industrial scenarios is a problem.

It is better to break down this complicated problem into a few questions:

- What is the most urgent industrial issue to be addressed?
- What is the most valuable and feasible solution to this problem?
- What is the most suitable technology for this solution?
- How to design a validation experiment?
- How to make rational use of certain techniques and tools to implement an experiment?

It has been at least six years since the beginning of Industry 4.0 or Industry Internet. There has been too much advertisement and boasting. Many digital industrial companies have claimed that they have helped customers create much value, but the real effects have not been shown in public. As a digital industrial leader, GE has spent billions of dollars on the development of digital technologies. However, GE sold its industrial solution division in September 2017 because of its unexpected outcomes. Venture capitals firms have made significant investment in the potential digital industrial startups. Numerous industrial IoT platforms are emerging in the market, but a number of them fade soon after because of a lack of data, domain expertise, and practical experience. Some people doubt about the future of the digital industry.

Many people have heard about industrial maintenance technology, but they have been confused about concepts such as reactive maintenance, planned maintenance, proactive maintenance, and predictive maintenance. What are problems in other industrial scenarios such as operations or quality control besides maintenance? Please see the figure 0 below for the Industrial IoT problem taxonomy raised by Hitachi.

IIoT Problem Taxonomy NEXT 2017

Analytics	Maintenance	Operations	Quality
Descriptive	1. Equipment Monitoring 2. Performance Analytics 3. Maintenance Analytics 4. Equipment Failure Root Cause Analysis	1. Operations Monitoring 2. Characterize Process 3. Operator Behavior 4. Operation Failure Root Cause Analysis	1. Quality Monitoring 2. Testing Process Monitoring & Evaluation 3. Detect Quality Loss 4. Defect Root Cause Analysis
Predictive	1. Predict Failures 2. Estimate RUL 3. Predict Failure Impact	1. Predict Activity Time 2. Predict Production KPI(s) 3. Demand Forecasting 4. Supply Chain Disruption	1. Early Defect Detection 2. Yield Quality Predict.
Prescriptive	1. Reduce Failure Cost 2. Reduce Failure Rate 3. Repair Recommendation 4. Optimize Maintenance	1. Failure Rate Reduction 2. Fuel/Energy Reduction 3. Equipment Scheduling and Dynamic Dispatch 4. Operations Recommendation	1. Process Parameter Recommendation for Quality Improvement 2. Improve Testing

© Hitachi, Ltd. 2017. All rights reserved.

Figure 1. Industrial IoT Problem Taxonomy (Source: [1])

Many people know that big data is essential in industrial sector, but they do not know what characteristics industrial data have, or how to process, analyze, and extract value from industrial data. Many people value AI and have high expectations for of AI's performance in industry, however, they do not know in what kind of industrial scenarios AI plays a valuable role, or what kind of AI methods are suitable for certain industrial problem, or how to use AI algorithms and tools to implement a research project.

1.2 Research Objectives

The primary objective of this research is to figure out the most valuable and urgent industrial scenario and issue as well as the most suitable technology that is able to solve the problem in this scenario. To investigate these industrial scenarios and issues, I overview and analyze the entire ecosystem of the digital industry. I then formulate a valuable and feasible solution to the problem in the scenario. After analysis of this solution, I design an experiment, implement it, and evaluate the result in order to make sure this solution is indeed able to optimize results and create value. Finally, I summarize not only the experiment but also the entire approach and find out where there is room for improvement in the future research.

The second objective is to demonstrate how to use AI technologies such as machine learning and deep learning to solve industrial problems step by step through hands-on practice in an experiment that particularly designed for this solution.

This research is a good opportunity for me to not only review knowledge and skills I learned in the core course of the System Design and Management program but also apply it in practice by using system thinking to analyze and solve a problem from industry.

In sum, all of the efforts would not only help people to better understand the ecosystem of digital industry and related valuable solutions but also benefit my future endeavors in the digital industry.

1.3 Research Approach

This investigation uses a system approach, including system architecting, system engineering, and project management, to complete the process of analysis, design, and implementation. This thesis describes the approach of this research as follows:

1. **Introduction:** The motivations and objectives of this research on how to figure out the industrial scenario, issue, and technology, and how to implement a solution to solve the problem.

2. **Ecosystem Analysis:** The analysis of the entire digital industrial ecosystem in various dimensions such as initiatives, technologies in related domains, stakeholders, markets, and strategies.

3. **Solution Analysis:** The analysis of the predictive maintenance solution in various dimensions such as background, importance, suitable scenarios, market, business model, and technology.

4. **Experiment Design:** The plan for the predictive maintenance solution, including goal, data source and description, methods, steps, and tools.

5. **Experiment Implementation:** The baseline approach and optimal approach to implement the experiment, including data preparation, selection and evaluation of both regression and classification models, deep learning practice through neural network building and optimization.

6. **Conclusion:** Summary, contributions and expectation, and limitations and future research

Chapter 2: Ecosystem Analysis

2.1 Background Analysis

"In the future, all the manufacturers make the machines, the machines can not only produce the products, the machines must have talked the machine, must have think, and the machine is not going to be supported by oil by electricity, the machine is going to be supported by data. In the future world, business will not focus on the size, business will not focus on standardization and the power, they will focus on the flexibility, nimbleness, customization, and user friendliness."
-- said by Jack Ma (founder and executive chairman of Alibaba Group) on the Hannover Messe 2015

These words represent future manufacturing in both German and US philosophies. The Germans call it smart manufacturing to meet requirement of customization, socialization, and flexibility, while Americans call it intelligent manufacturing to highlight the intelligence of physical system [2].

2.1.1 Industry 4.0

Industry 4.0 is a name for the current trend of automation and data exchange in manufacturing technologies. The term "Industrie 4.0" originates from a project in the high-tech strategy of the German government, which promotes the computerization of manufacturing. The term "Industrie 4.0" was revived in 2011 at the Hannover Fair [3]. The roadmap of industrial revolutions is shown in the figure 2 below.

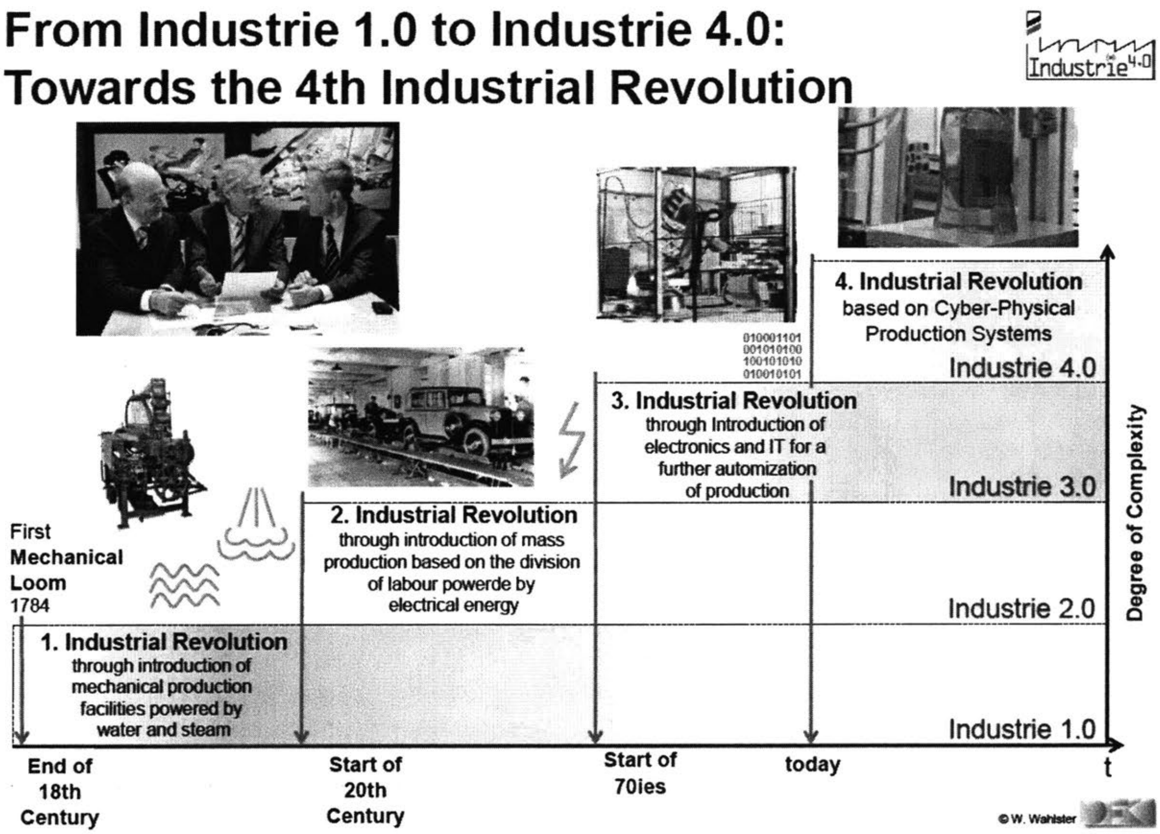

Figure 2. Roadmap of Industrial Revolutions (Source: [4])

Industry 4.0 focuses on cyber-physical production systems for mass customization. With the cyber-physical system, Industry 4.0 promotes digitalization and smartness of information of supply, manufacturing, and sales, achieves fast, effective, and customized supply of products for a reasonable price based on semantic technologies and service matchmaking, and enables Plug & Produce and Multi adaptive Smart Factories. In Industry 4.0 scenario, plant workers are assisted by collaborative robots, intelligent industrial systems, augmented reality devices, etc [4].

Industry 4.0 is a success story of a strategic public-private partnership and secures Germany's economic power as a leader in manufacturing. Typical industry giants in Germany are Siemens, SAP, Bosch, etc. Industry 4.0 is the foundation of digital economy.

Four design principles in Industry 4.0 are shown in the figure 3 below.

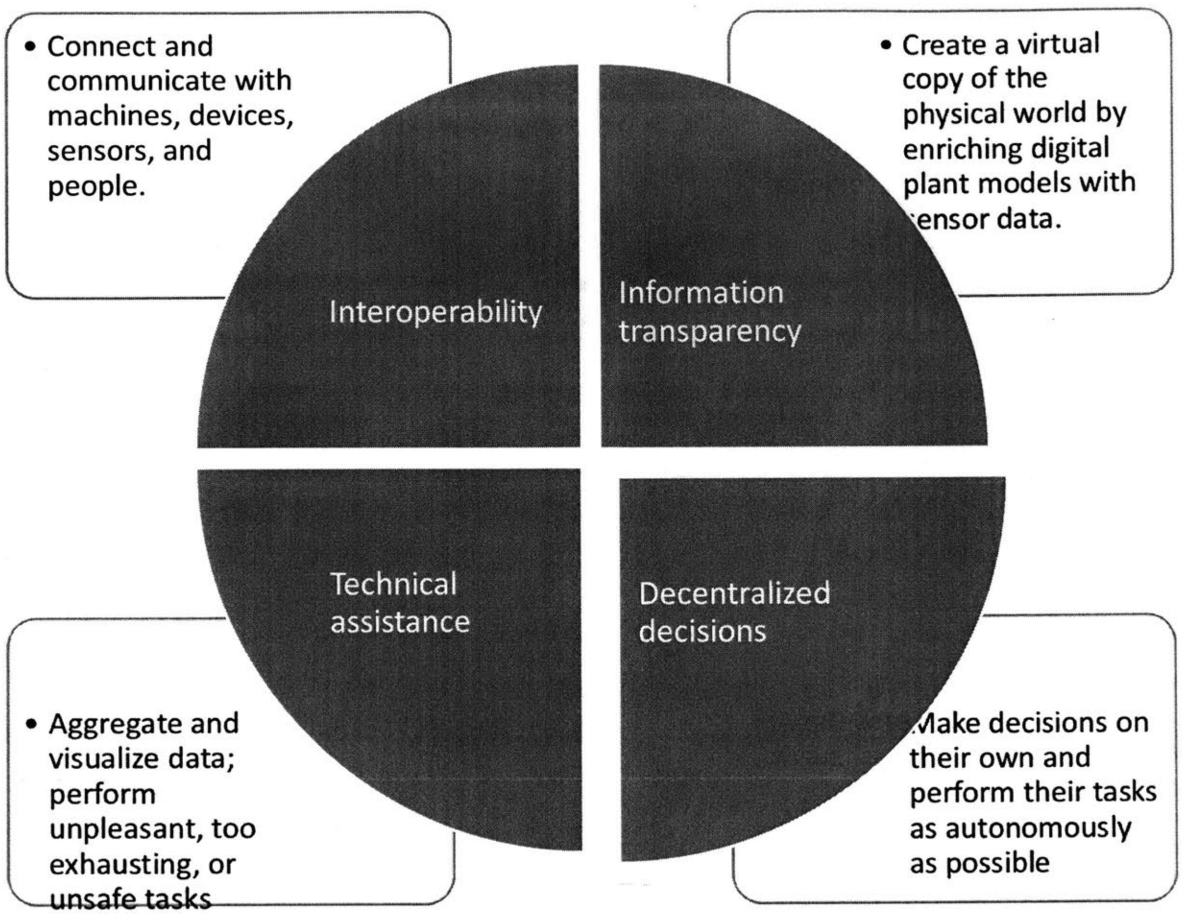

Figure 3. Design Principles in Industry 4.0

2.1.2 Industrial Internet

As an industry giant in US, GE promoted digital industry and woke up as a software and analytics company. GE proposed the Industrial Internet concept in late 2012 through the whitepaper – Industrial Internet: Pushing the Boundaries of Minds and Machines by Peter C. Evans and Marco Annunziata. Industrial Internet was referred to as Industrial IoT or IIoT later. Based on GE's estimation, the Industrial Internet could be a $225 billion market by 2020. GE Digital was built to explore digital transformation, and lay a solid digital foundation in GE. With significant investments and resources in the Industrial Internet, GE is driving its own digital industrial transformation. With its experience and expertise accumulated during its own transformation, GE is helping customers achieve their digital transformation [5]. Other typical industry companies in US are Honeywell, Emerson, Rockwell, etc.

As a “network of networks”, the “regular” Internet networks people with information, while the Industrial Internet connects machines, devices, systems, plants, industries and people to collect, process, analyze industrial customers’ data, and create value through data-driven digital technologies [6].

Through a convergence of Operation Technology (OT) and Information Technology (IT), the Industrial Internet is weaving together OT systems with IT systems, and achieving end-to-end automation solution for complex processes: production monitoring, repair and maintenance, and asset management and optimization.

As one of the members, GE co-founded the Industrial Internet Consortium (IIC) to accelerate the development, adoption, and widespread use of Industrial Internet, and create value from connected and intelligent machines, devices, systems, and people at work.

To achieve consistency of Industrial Internet systems, the Industrial Internet Consortium (IIC) proposed the Industrial Internet Architecture Framework and Industrial Internet Reference Architecture. Based on the architecture, Industrial Internet system has become a distinct domain, similar to control, operations, information, application, and business [7]. Please see the figure 4 below for the overview of functional domains and the figure 5 below for the mapping between a three-tier architecture to the functional domains.

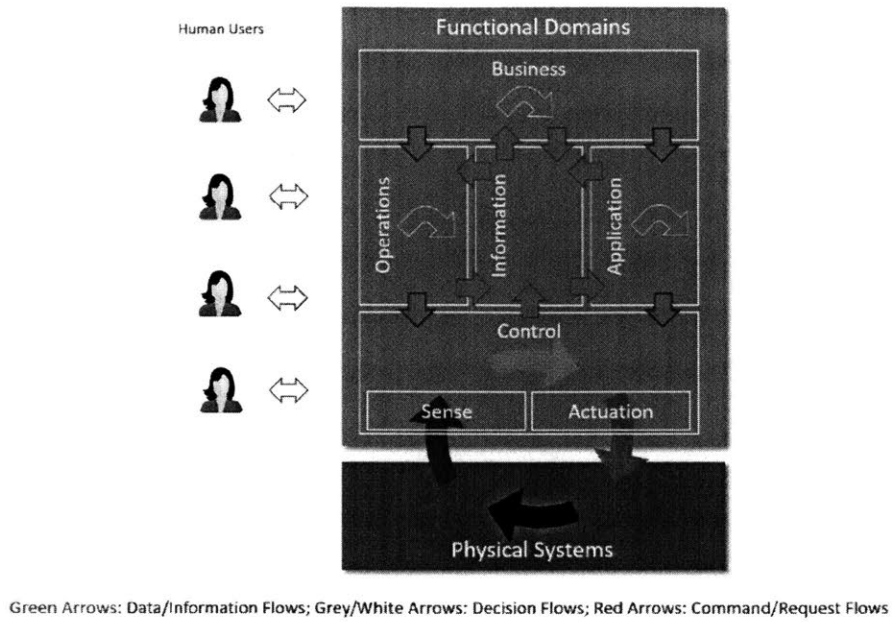

Figure 4. Functional Domains (Source: [7])

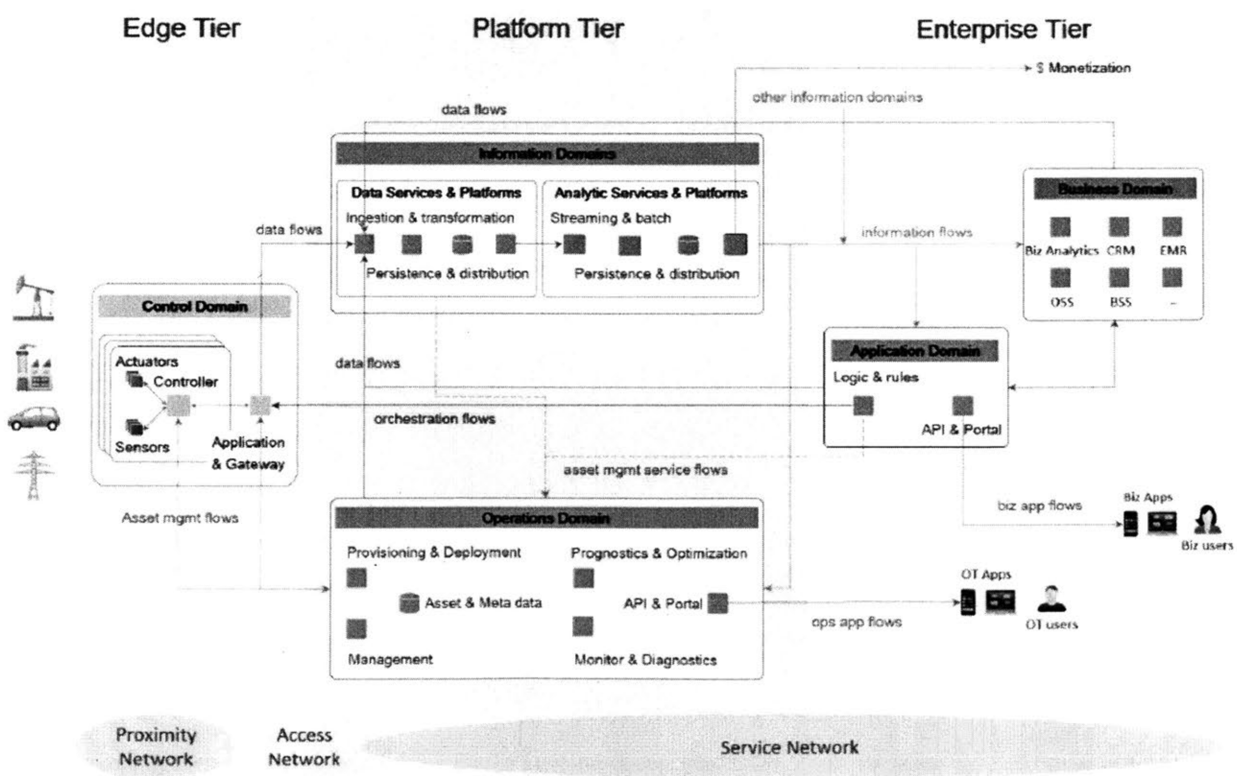

Figure 5. Mapping between a Three-tier Architecture to the Functional Domains (Source: [7])

2.1.3 Made in China 2025

In 2015, the Chinese government established the "Made in China 2025" initiative to upgrade China from a manufacturer of quantity to that of quality.

"Made in China 2025" proposed that persist in the basic principle of "innovation-driven, quality first, green development, structural optimization, and talent-based", and persist in the basic principle of "market-oriented, government guided, current status-based, long-term targeted, comprehensive progressive, and breakthrough at key points, independent development, and open and collaborative". Through "three steps" to achieve the strategic goal of being a strong manufacturing country [8].

"Made in China 2025" clearly defined 9 strategic tasks, prioritized them, and proposed 8 strategic supports and guarantees. This timeline of this goal is shown in the figure 6 below, and 10 key areas that promote breakthroughs are shown in the figure 7 below. Typical industry companies in China are Huawei, SANY, Foxconn, etc. In addition, BAT (Baidu, Alibaba, Tencent) are seizing this great opportunity to enter this market.

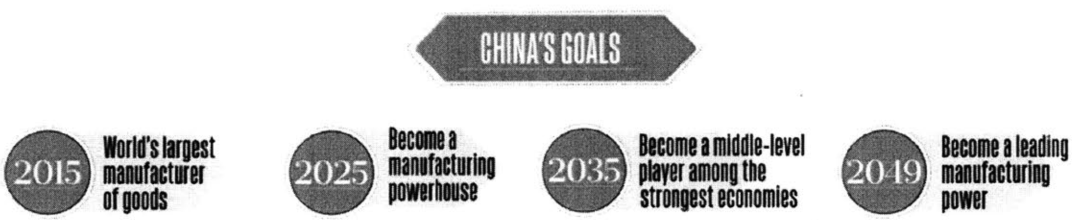

Figure 6. Goals of Made in China 2025 (Source: [9])

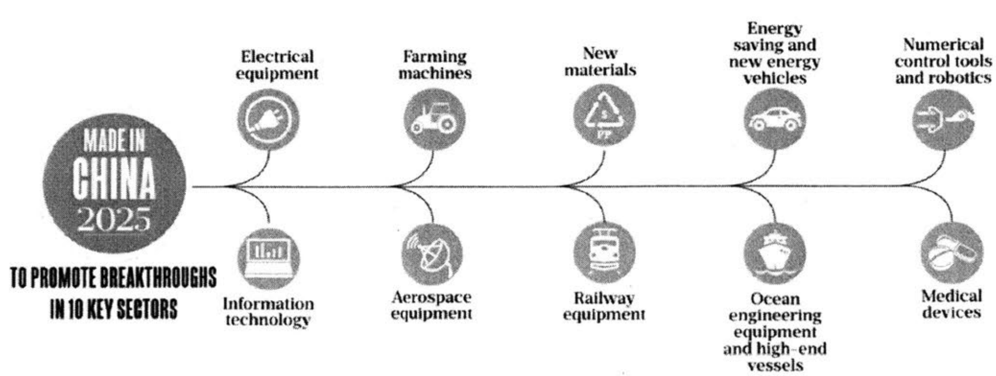

Figure 7. 10 Key Areas that Promote Breakthroughs (Source: [9])

2.1.4 The Industrial Value Chain Initiative (IVI)

Founded in Japan, IVI is a forum to design a new society through combining manufacturing and information technologies, and for all enterprises to collaboratively take an initiative. Typical industry companies in Japan are Hitachi, Mitsubishi, Yokogawa, etc.

Traditionally, Japanese industrial companies make everything they need by themselves. This is not sustainable in the globalized world today. IVI aims at creating a mutually connected system architecture for a cooperation among companies, in particularly, small and mid-sized enterprises (SME's). IVI offers opportunities for correlated companies to leverage their own strengths and advantages to interact and work with each other in order to scale up their business size and explore emerging market in the digital industry.

Connected Manufacturing and Loose Standards are two principles of IVI.

- Connected Manufacturing: prevent from overburden, waste, and unevenness through connecting plants and enterprises; create smart value chains based on industrial automation and human ability simultaneously.
- Loose Standards: use an adaptable model instead of a rigid system. A strict standard faces challenges in complex manufacturing settings with old and new elements, while a loose standard in the reference model enables interconnection case-by-case [10]. Please see the figure 8 below for the loose standard in the reference model.

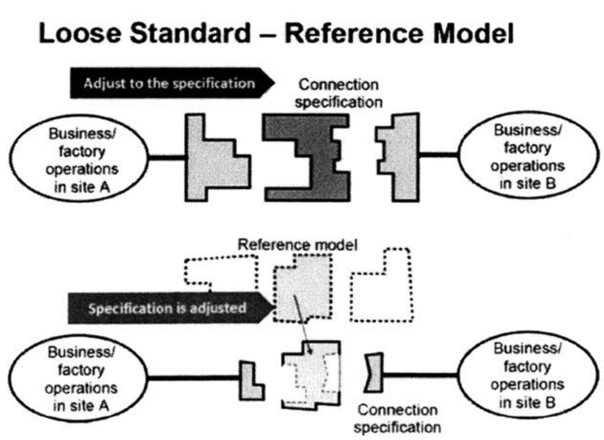

Figure 8. Loose Standard in the Reference Model in IVI (Source: [10])

2.1.5 Alternatives Initiatives

Many European countries joined the German industry 4.0 initiative, to name a few: Sweden, Austria, Ireland, France, etc. The representative industrial companies are ABB, Schneider, and the like. Through the Horizon 2020 program, the European Commission is promoting the development and adoption of digital technologies in European countries to reshape their industries. Enormous progress is witnessed in 3D printing, additive manufacturing, IoT, and robotics in Europe in these years.

In Asia-Pacific region, besides China and Japan, many other Asian countries such as India, South Korea, Singapore are taking initiatives to motivate implementation of the digital industry. Besides major company like Samsung, some potential startups such as Flutura and Altizon are coming into our sight.

I use the term "digital industry" to represent all alternatives of Industry 4.0 in the following paragraphs.

2.2 Technology Analysis

The digital industry is driven by technology. Key technologies contribute to the digital industry include but no limited to [6] [11]:

- Sensors and actuators
- Robotics
- M2M and machine protocols
- Network and IoT
- Control systems, SCADAs, DCSs, PLCs
- Data management and data analytics
- AI and machine learning
- Network and connectivity
- Cloud computing

- Edge computing
- Mobile and Wearables
- HMI, UI/UX
- Augmented Reality (AR) & Virtual Reality (VR)
- Digital Twins
- Cybersecurity
- Block Chain

Like sensor and industrial control system (ICS), traditional industrial technologies lay a solid foundation for the digital industry. Digital technology is a new power that transforms the industrial sector throughout the world. Based on BCG's research, nine technology trends form the building blocks of Industry 4.0, showing in the figure 9 below. In addition, a digital compass in a McKinsey report illustrated digital technologies and their value propositions. Please see the figure 10 below for the digital compass. I would like to introduce key technologies that exert significant impact in the digital industry.

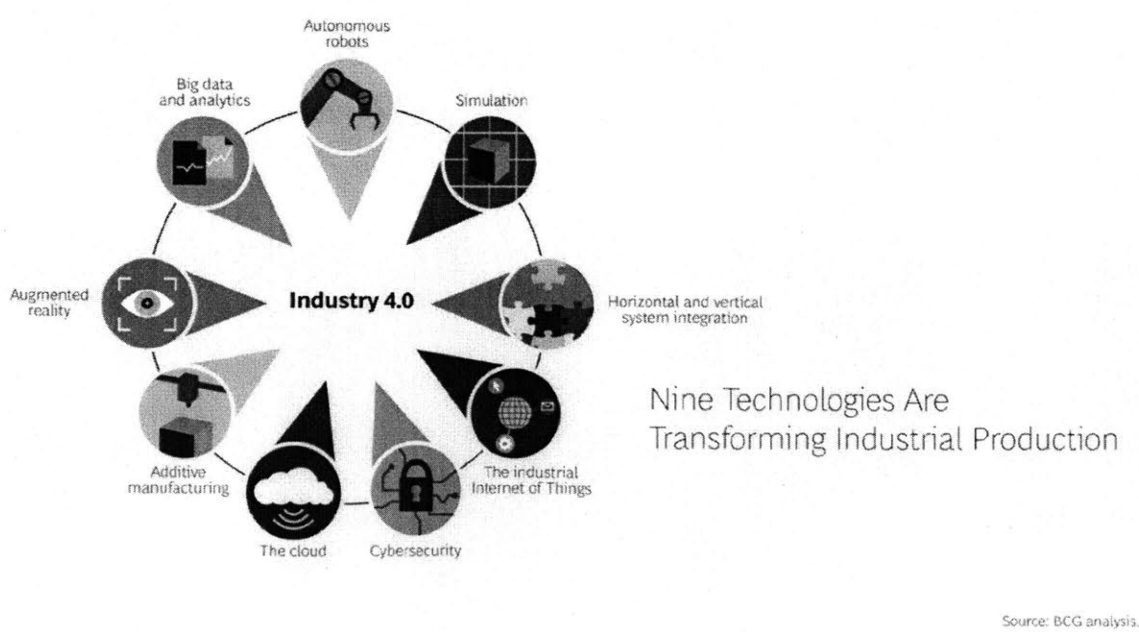

Figure 9. Technologies that Transform Industrial Production (Source: [12])

Figure 10. Digital Compass (Source: [13])

2.2.1 IT/OT Convergence

As is known to all, Information Technology (IT) needs not any words here. I would like to introduce OT here.

Operational Technology (OT) – the hardware and software dedicated to detecting or causing changes in physical processes through direct monitoring and/or control of physical devices such as valves, pumps, etc. Simply put, OT is the use of computers to monitor or alter the physical state of a system, such as the control system for a power station or the control network for a rail system. The term has become established to demonstrate the technological and functional differences between traditional IT systems and Industrial Control Systems environment.

-- Wikipedia [14]

Here is an analogy. The Internet is a central nervous system, the cloud acts as a brain, while OT components make up the body, giving the Internet eyes and ears, arms and fingers to gather information and act upon that information. OT components have capabilities for local automation and execution.

Good examples of OT are Industrial Control Systems (ICSs) such as Distributed Control Systems (DCS), Programmable Logic Controllers (PLC), Supervisory Control and Data Acquisition (SCADA) Systems, and Remote Terminal Unit (RTU). ICSs are used to monitor and control the processes and interactions among sensors and actuators.

ICSs process operational data from electronic devices, telecommunications, computer systems, and monitor various process values, such as temperature, pressure, flow, level, etc. ICSs also control engines, conveyors, pumps, valves, fans and other machines and equipment to regulate corresponding process values to prevent them from dangerous conditions. ICSs process real time or near-real time data with high requirements of availability and reliability.

OT-standard industrial communications protocols are Modbus, Profibus, etc. Gradually, IT-standard network protocols such as TCP/IP are being adopted in OT components in order to reduce complexity and increase compatibility, but the tradeoff is the reduction in security for OT systems.

ICSs have existed for as long as industrial processes. ICSs take on new meaning with digital technologies, in particularly, IT/OT convergence. Based on Gartner's prediction, by 2020, 50 percent of OT service providers will create key partnerships with IT-centric providers for IoT offerings." [15]

IT/OT convergence is reshaping long-standing processes in almost every industry to enable complex systems to monitor, maintain, control and optimize themselves, and remove the necessity for human involvement in many industrial scenarios. Through this convergence, best practices in IT such as software development, deployment and operations are being adopted in software-defined OT systems. In addition, IT systems such as Enterprise Resource Planning

(ERP) and Asset Management System are being seamlessly interfaced with OT systems (please see the figure 11 below), facilitating end-to-end automation solution and application such as predictive maintenance [16]. I will take a deep dive into it in the following chapters.

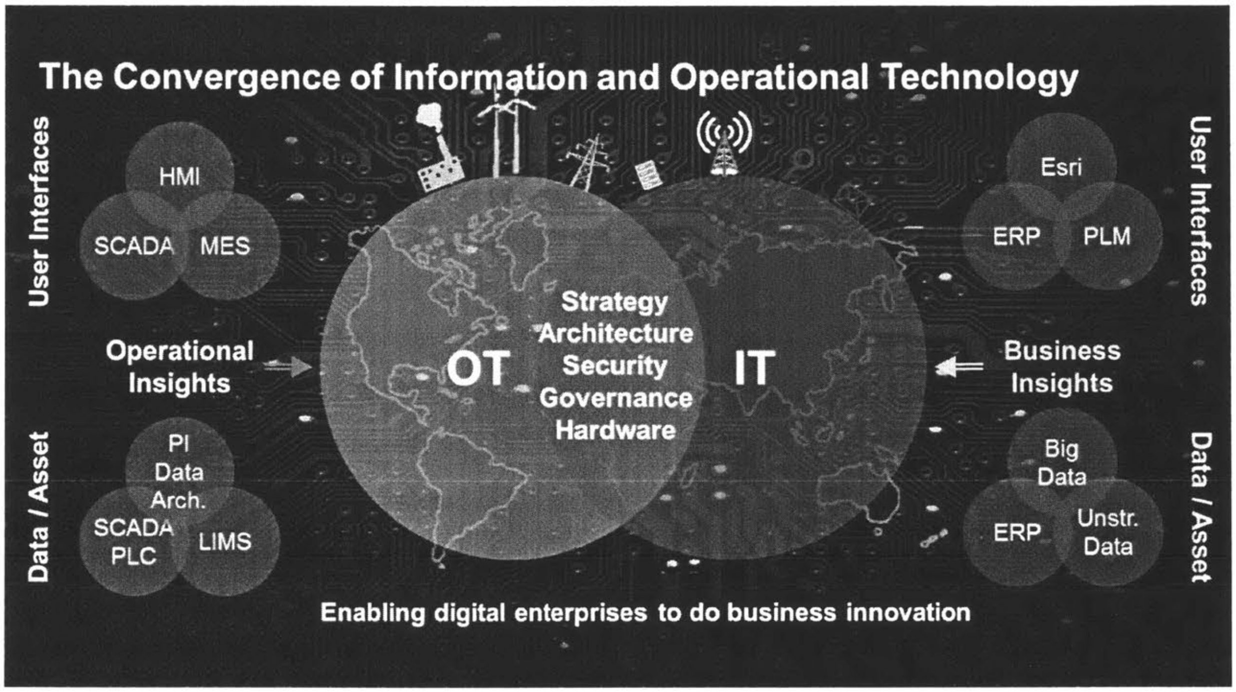

Figure 11. Convergence of IT/OT Systems (Source: [17])

2.2.2 M2M and Machine Protocols

Machine-to-machine refers to direct communication between devices using any communications channel, including wired and wireless. Machine to machine communication can include industrial instrumentation, enabling a sensor or meter to communicate the data it records (such as temperature, inventory level, etc.) to application software that can use it (for example, adjusting an industrial process based on temperature or placing orders to replenish inventory). Such communication was originally accomplished by having a remote network of machines relay information back to a central hub for analysis, which would then be rerouted into a system like a personal computer.

--Wikipedia [18]

Machine-to-machine (M2M) communication represents technological solutions and deployments enabling machines, devices and objects to communicate with each other without human interactions.

M2M technology is a basis of the digital industry. It enables the actual connection and interaction among machines, either directly point-to-point or over the Internet, serial, Ethernet, or other local LANs under specific protocols.

Compared with Web communication, OT communications are more complicate because they interact with system-external environments and they need high speed data transmission. Please see the table 1 below for protocols currently being used in M2M and figure 12 below for protocols in OSI model. No one protocol seems to become a standard any time soon.

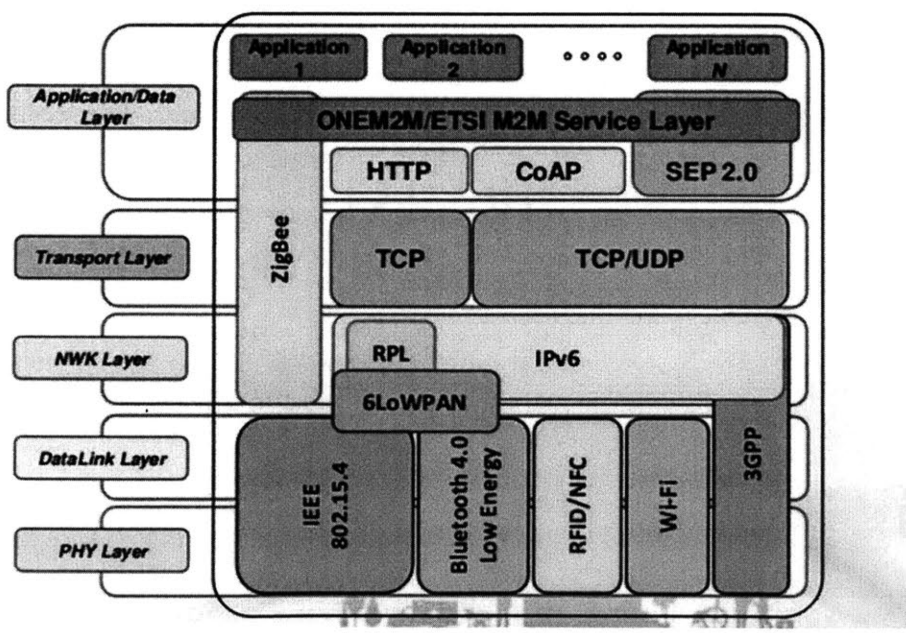

Figure 12. M2M Protocols in OSI Model (Source: [19])

Table 1. Protocols for M2M (Source: [20])

PROTOCOL	DESCRIPTION
MQTT	A publish-subscribe protocol used over TCP/IP. Lightweight, low code footprint, minimal bandwidth.
CoAP	Constrained Application Protocol Application layer protocol used for constrained (low-power, low-memory, etc.) nodes and networks.
AMQP	Advanced Message Queuing Protocol Application layer, wire-level protocol that supports a variety of messaging patterns.
HTTP/2	Updated version of Hypertext Transfer Protocol Built with HTTP 1.1 compatibility and performance enhancement in mind.
IPv6	Internet Protocol Version 6 Updated version of the Internet Protocol Version 4, necessary for assigning unique addresses to the rapidly growing number of machines connected to the Internet (due partially to the increase of Things and M2M connections).
6LoWPAN	IPv6 over Low power Wireless Personal Area Networks The 6LoWPAN group has defined encapsulation and header compression mechanisms that allow IPv6 packets to be sent and received over IEEE 802.15.4 based networks.

2.2.3 Convergence of H2H, H2M, and M2M

Besides convergence of IT and OT, we also value the convergence of Human-to-Human (H2H), Human-to-Machine (H2M), and Machine-to-Machine (M2M) communication.

In the past, H2H and H2M technology was widely used. Without machine, human (operators, engineers, managers) have to communicate with each other and play their own roles in production process. With the development of ICSs, basic semi-automatic production was realized through operation on the Human-machine interface (HMI). Please see the table 2 below for the form and function of H2H, H2M, and M2M.

Table 2. Form and Function of H2H, H2M, and M2M

Form	**Function**
Convergence of H2H, H2M, and M2M: Different kinds of machines, such as on-site instruments and devices, operation systems, and information systems, have been connected to each other through network. Human such as operators, engineers and managers, are also involved in the network.	M2M: Machines are able to communicate with each other to exchange data, complete interlock control tasks.
	H2M: Machines are able to send significant information to human in important situation, for instance, turbines can inform engineers for overhaul through sending message to the mobile operating device; Human can obtain all of information they need to control production process.
	H2H: Based on all of information they get from machines, Human are able to interact with each other and better improve their work efficiency.

With the rapid development of M2M technology, the convergence of H2H, H2M, and M2M becomes reality. I use a Model-based System Engineering (MBSE) tool - Object-Process Methodology (OPM) to illustrate this convergence. Through the industrial internet, devices, plant, technologies (IT & OT), and people are closely connected, and the convergence of H2H, H2M, and M2M is realized. Please see the figure 13 below for the OPM chart.

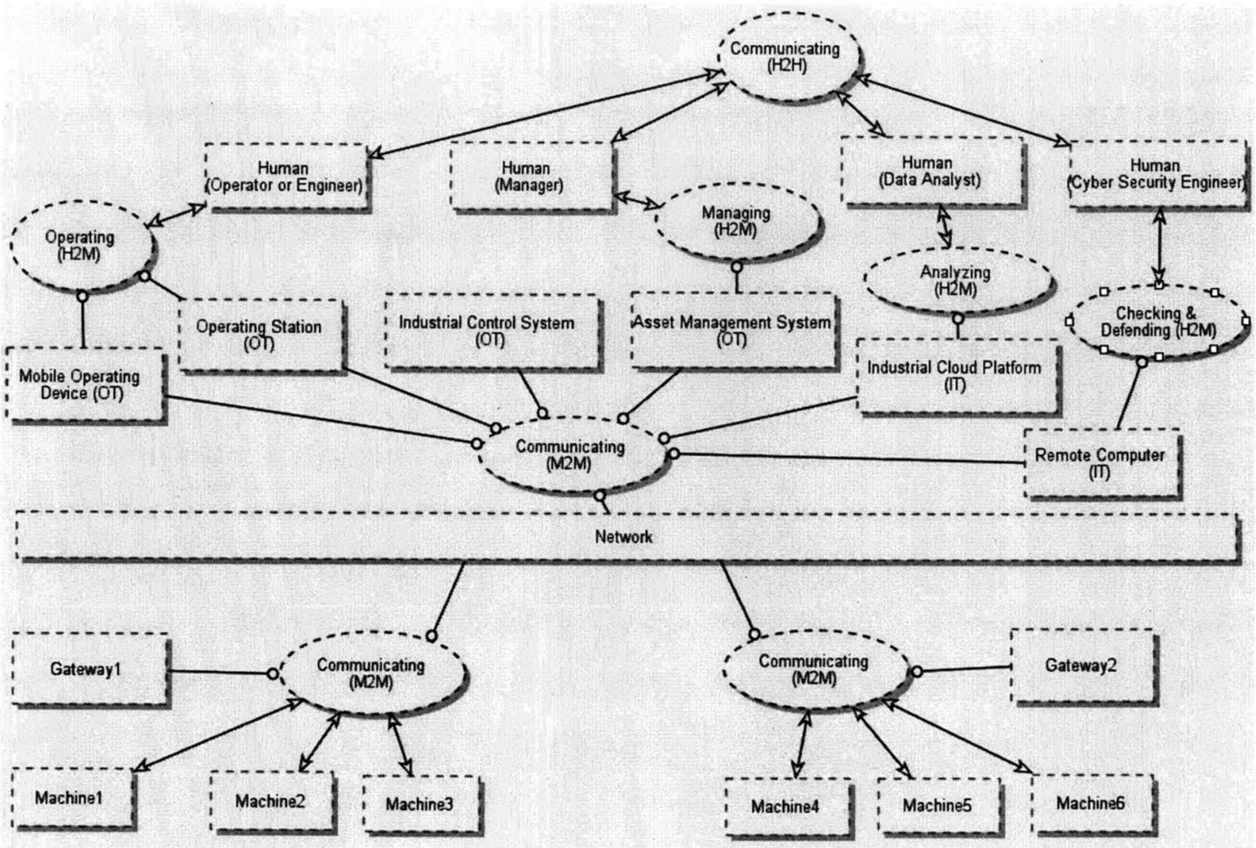

Figure 13. Convergence of H2H, H2M, and M2M

2.2.4 Internet of Things (IoT)

The Internet of Things (IoT) is the network of physical devices, vehicles, home appliances and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these objects to connect and exchange data. Each thing is uniquely identifiable through its embedded computing system but is able to inter-operate within the existing Internet infrastructure.

--Wikipedia [21]

M2M provides IoT with the connectivity that enables capabilities, while IoT has a horizontal approach that polls vertical applications together; M2M focuses on direct point-to-point connectivity across mobile or fixed networks, while IoT enables communications with IP networks and cloud platforms; M2M focuses the communication only, while IoT has broader processes and applications; In sum, M2M applications are subset of IoT infrastructure. IoT provides context for data and events across applications, groups and organizations. IoT goes beyond M2M and provide and utilize its extensive resources [22]. The difference between M2M and IoT is simply shown in the figure 14 below.

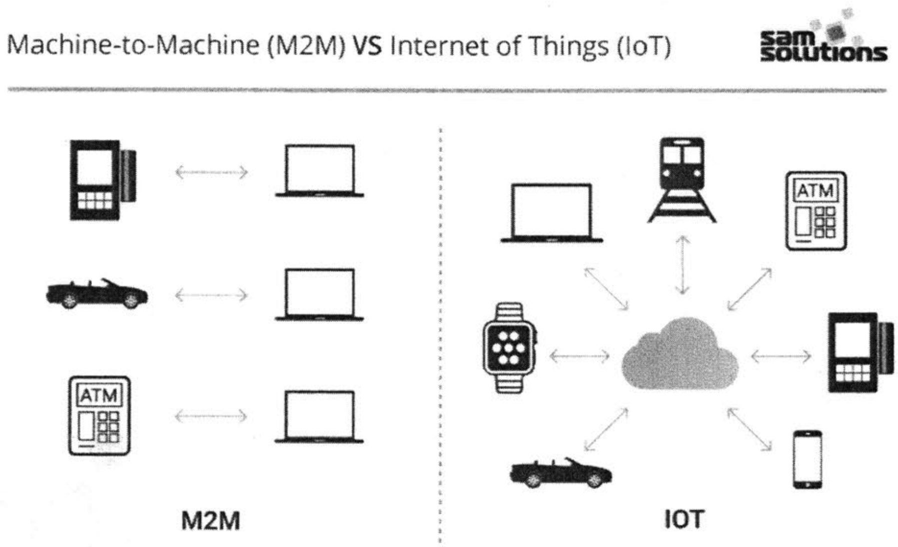

Figure 14. Difference between M2M and IoT (Source: [22])

In a broad sense, industrial IoT equals to the initiative - industrial internet, while in a narrow sense, it refers to IoT hardware and software application in industrial sector. We discussed the broad sense in the 2.1.2 sector. Here we discuss it in the narrow context.

Based on a report by Grand View Research, Inc., the global industrial IoT market is expected to reach USD 933.62 billion by 2025 [23].

Narrow-Band (NB) IoT, LoRa (from long range), LTE CAT-M are adopted in the digital industry. Comparison of various IoT network is shown in the table 3 below.

Table 3. Comparison of Various IoT Network (Source: [24])

basvankaam.com sharing knowledge

The Internet of Things networking technology cheat sheet 1.0

Network:	Sigfox	LoRa	NB-IoT (Cat NB1)	LTE-M (Cat M1)	LTE Cat 0	LTE Cat 1
Type:	PLWAN	PLWAN	DSSS modulation	LTE (cellular)	LTE (cellular)	LTE (cellular)
Low Power:	+++++	++++	++++	+++	++	++
Throughput Kbit/s:	0,1	50	100	375	1000	10.000
Bandwidth:	Ultra-narrowband	Narrowband	Narrowband	Low	High	High
Latency:	1 – 30s	Based on profile	1.6 – 10s	10 – 15ms	Unknown	50 – 100ms
Standard:	Proprietary	Proprietary	3GPP Rel. 13	3GPP Rel. 13	3GPP Rel. 12	3GPP Rel. 8
Availability world-wide:	++	+++	++	++	+++++	++++
Spectrum:	Unlicensed ISM	Unlicensed ISM	Licensed LTE	Licensed LTE	Licensed LTE	Licensed LTE
Complexity:	Very low	Low	Very low	Low / medium	High	High
Coverage / range:	Medium / high	Medium / high	High	High	High	High
Battery life:	Very high	Very high / high	High	Medium / high	Low	Low
Gateway needed:	Yes	Yes	No, but optional	Optional	Optional	Optional
Signal penetration:	High	Medium / high	Medium / high	Medium / high	Low	Low
Security:	+++	+++	+++	+++++	++++	++++
Future proof:	+++	+++	+++++	+++++	+++	+++

See the accompanying blog series on basvankaam.com for more details on some of the abovementioned features/characteristics

Saturday, July 15, 2017 - Twitter @BasvanKaam

A series of Industrial IoT technology drives operational efficiency in many cases in the digital industry. For example, using Narrow Band IoT or LoRa technology to bring environment monitoring sensor data into control system or cloud platform, and evaluate the impact of production on the environment; monitoring real-time data from crew well-being, and identifying dangerous situations for personnel safety management.

2.2.5 Network and Standard

Classic automation pyramid is built with layers of network (Machine, Control, Manufacturing, Enterprise) with corresponding communication protocols and standards (Fieldbus, Industrial Ethernet, Ethernet, Internet). Through levels of monitoring, control, and management, information flows upwards from field devices to enterprise. The communication is not as smooth as we expected, and issues regarding compatibility between layers frequently come around because of not only diversity, customization, localization of industrial products but also various industrial communication protocols and standards.

Most of industrial automation (OT) players developed and promoted their own protocols and standards, such as Siemens's Profinet, Rockwell's ControlNet, DeviceNet, Ethernet/IP, Schneider's Modbus and TCP/IP, etc. We expect a standard that bridges the gap among various OT and IT products. Recently, a growing number of OT companies have adopted Ethernet as a standardized protocol in their OT products. Industrial Ethernet has been widely accepted in industries, it is more likely to enable the industrial Internet.

Please see the figure 15 below for the timescales for industrial communication standards. Industrial Ethernet like Profinet is concerned with sub-second timeframes. Stepping back from the sub-second timeframe to a longer time frame, industrial Ethernet becomes industrial Internet. This transformation happens as the granular sub-second data turns into information when analyzed over a longer time frame [25].

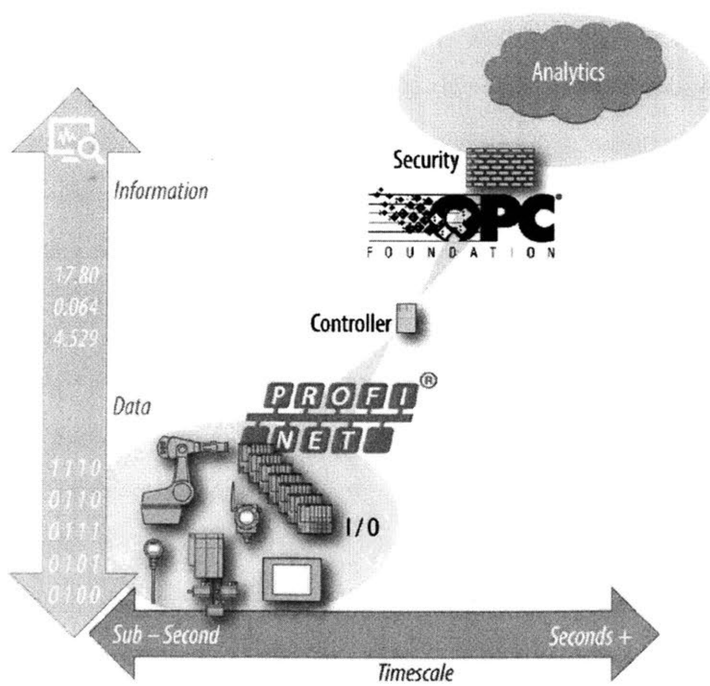

Figure 15. Timescale of Industrial Communication Standards (Source: [25])

Based on a report on Innovation Post, in 2017, for the first time, the market share of Industrial Ethernet exceeded Fieldbus. Please see the figure 16 below.

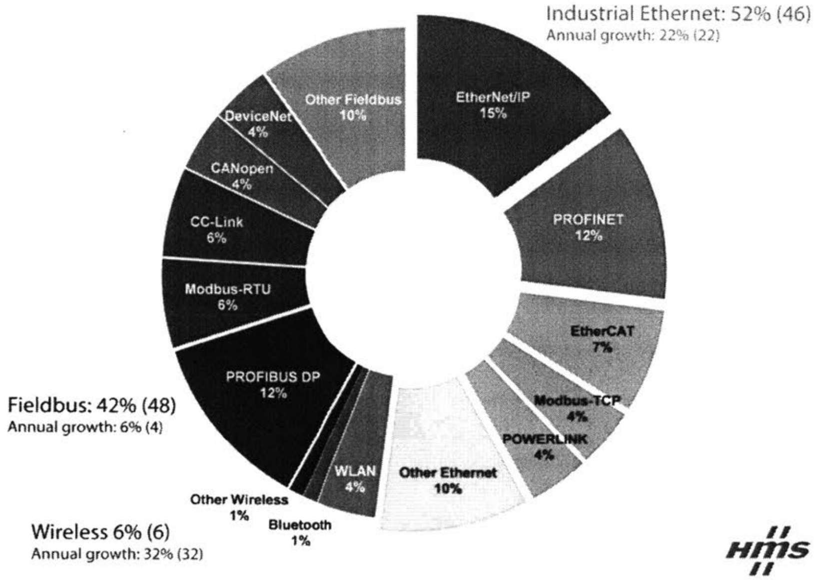

Figure 16. Industrial Communication Protocols and Standards (Source: [26])

As is known to all, Ethernet technology works in asynchronous model and solve issues regarding data link and network infrastructure share. However, any device can send data at any time in the network, and data transmission time is uncertain and inaccurate. Therefore, a real-time, certain, and reliable data transmission vehicle is expected. Now TSN and OPC UA catches our eyes.

2.2.5.1 Time-Sensitive Networking (TSN)

TSN is a set of standards under development by the Time-Sensitive Networking task group of the IEEE 802.1 working group. The standards define mechanisms for the time-sensitive transmission of data over Ethernet networks.

--Wikipedia [27]

TSN Add real-time functionality to IEEE 802 Ethernet. TSN is moving past the concept stage for industrial automation use. As is shown in the figure 17 below, TSN mechanism works in layer 2 in the Open System Interconnection (OSI) model. The characteristics of TSN is shown in the figure 18 below.

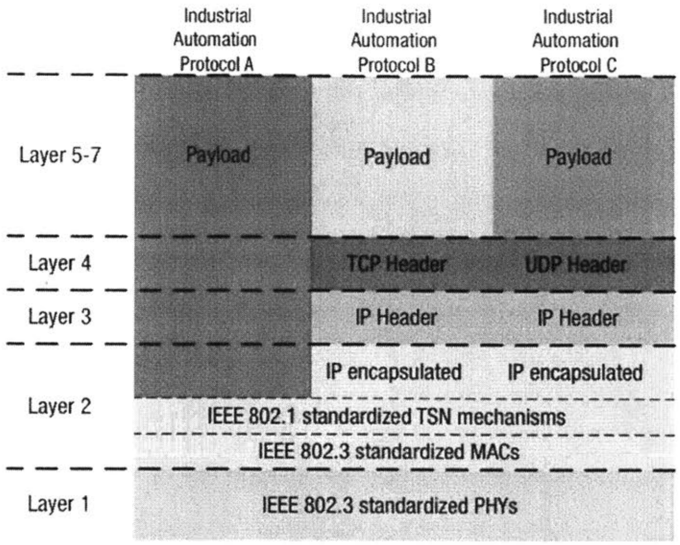

Figure 17. Layer for TSN (Source: [28])

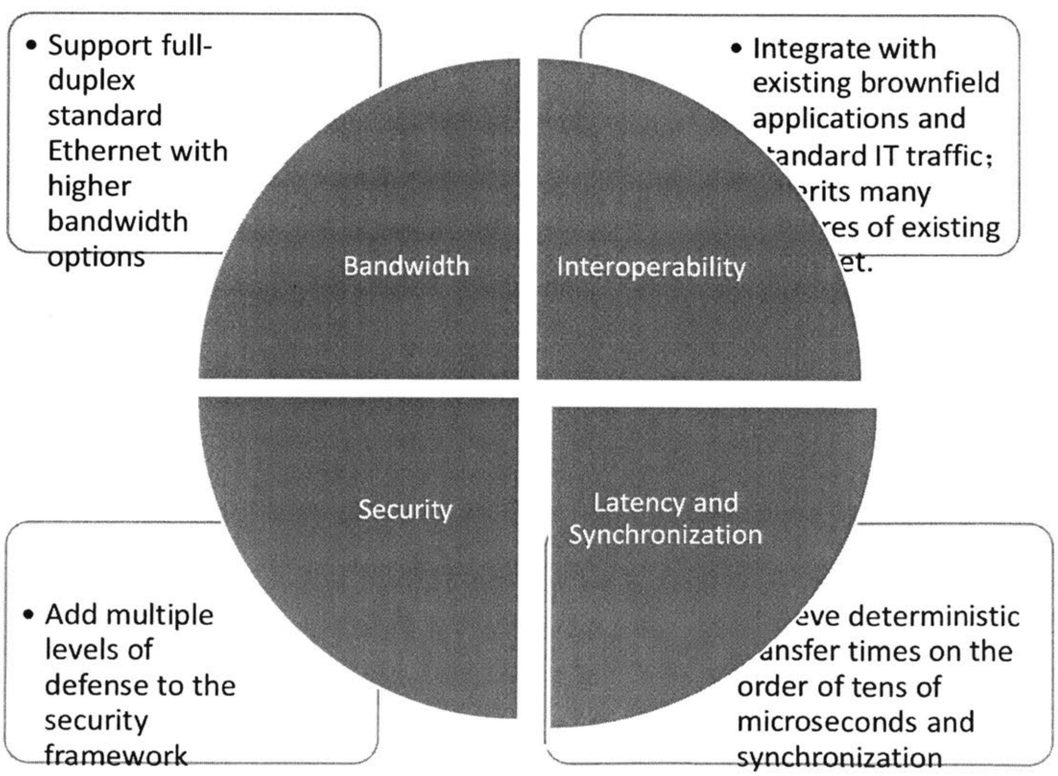

Figure 18. TSN Characteristics

TSN is enabling all data, including real-time information, to be transmitted through a single network in effect simultaneously over a shared network. In the deterministic network, our expectation is as follows [29] [30]:

- All clocks on all nodes are synchronized to a uniform network time.
- Time-critical data can be assured to be transmitted within a guaranteed amount of time, while other non-time-restricted data can be sent as normal.

2.2.5.2 OPC UA

Object Linking and Embedding for Process Control (OPC) specifies the industrial communication of real-time factory data between control devices and systems from different manufacturers.

OPC Unified Architecture (OPC UA) is a machine to M2M protocol for industrial automation developed by the OPC Foundation. It focuses on communication with industrial devices and systems for data collection and control.
--Wikipedia [31]

2.2.5.3 TSN + OPC UA

With the exponentially growing amount of data from machines and sensors, more effective and efficient network architectures are required. The conjunction of TSN and OPC UA catches our eyes in the digital industry recently. This solution meets the requirements for real-time, vendor-neutral Ethernet communication among machines, equipment, and systems in the digital industry. TSN deals with data acquisition problem, while OPC UA handles semantic parse issue.

This new architecture flattens the automation pyramid. Please see the figure 19 below the new structure. This provides a model where OPC UA 'clients' at management or enterprise levels can request data directly from OPC UA 'servers' at the device layer. From a new perspective, the automation pyramid is transforming to automation pillar. Please see the figure 20 below for the details in this transformation.

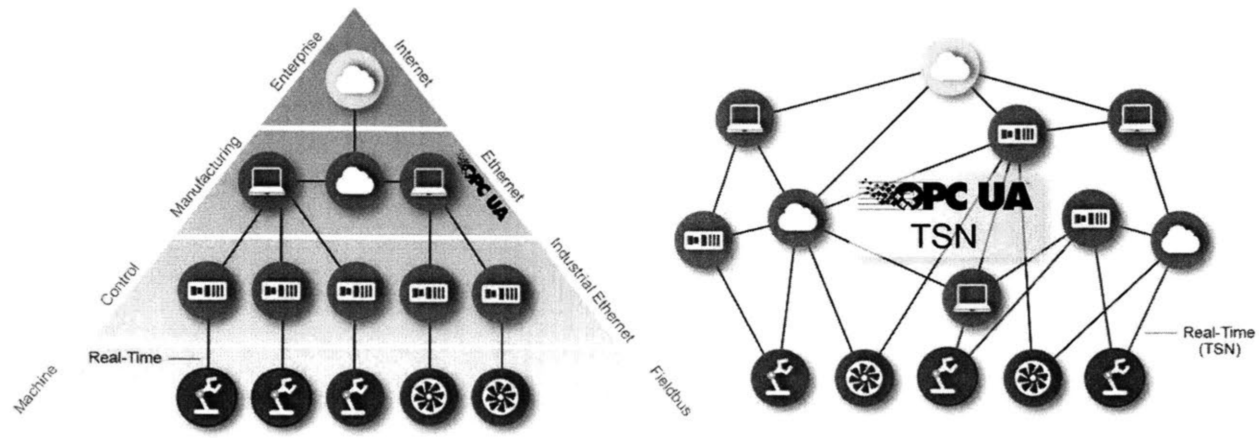

Figure 19. New Architecture that Flattens the Automation Pyramid (Source: [32])

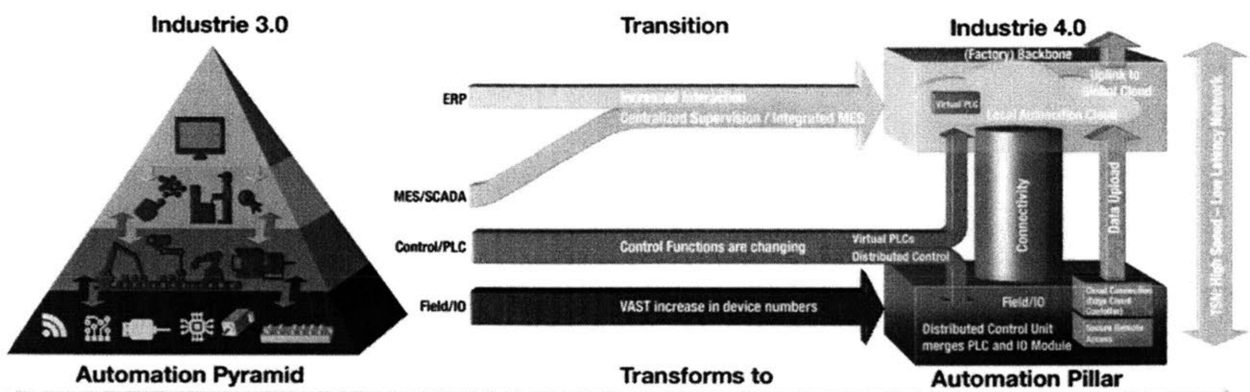

Figure 20. Transition of Architectures for the Automation (Source: [33])

OPC UA bridges certain gaps between OT and IT. A Publish/Subscribe model for OPC UA and TSN extension solve the problem of the proprietary and isolated networks used for real-time and safety sub-systems. This facilitates the exchange of real-time data from one-to-many and many-to-many over standard Ethernet.

Implementation of OPC UA TSN in the digital industry will take time. Actually, many industrial companies such as Honeywell and Rockwell joined OPC UA TSN initiative to facilitate the accelerated development of OPC UA TSN.

In addition, with the development of OPC UA TSN, more players outside of industrial automation, especially technology companies such as BAT in China, might enter the digital industry market.

2.2.6 Cloud Computing and Platform

Cloud computing is an information technology (IT) paradigm that enables ubiquitous access to shared pools of configurable system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet. Cloud computing relies on sharing of resources to achieve coherence and economies of scale, similar to a public utility.
-- Wikipedia [34]

Cloud computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing.
--Amazon Web Service [35]

Cloud computing means a lot in the digital industry. Because of its optimized architecture and super strong computing power, most industrial internet platforms are built on cloud in different models such as IaaS, PaaS, SaaS, etc.

Predix, the first cloud-based industrial internet platform, was developed and launched by GE for building and operating industrial applications. This Platform-as-a-Service (PaaS) platform enables fast deployment and elastic scale of cloud applications. Predix facilitates embedded software for standardized IoT connectivity, provides services and micro-services, manages huge amount of industrial data, analyzes industrial assets and processes for enterprise decision-making, and creates secured environments for applications' compliant management and operation at industrial scale.

The architecture of various services in different levels are shown in the figure 21 below. As a comprehensive platform, Predix works with components from edge to cloud in various use cases. The architecture of Predix is shown in the figure 22 below.

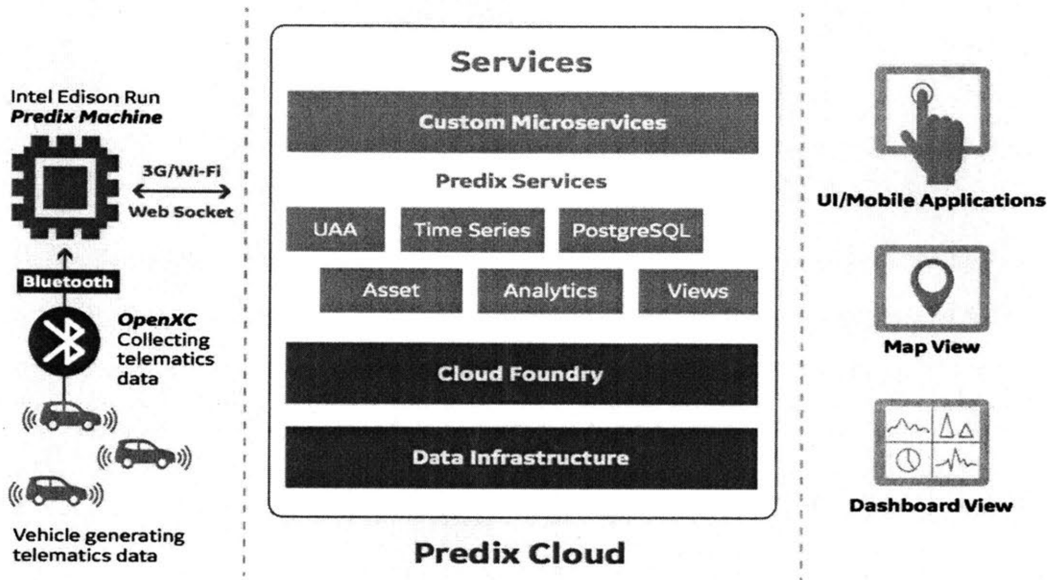

Figure 21. The Architecture of Various Services in Predix (Source: [36])

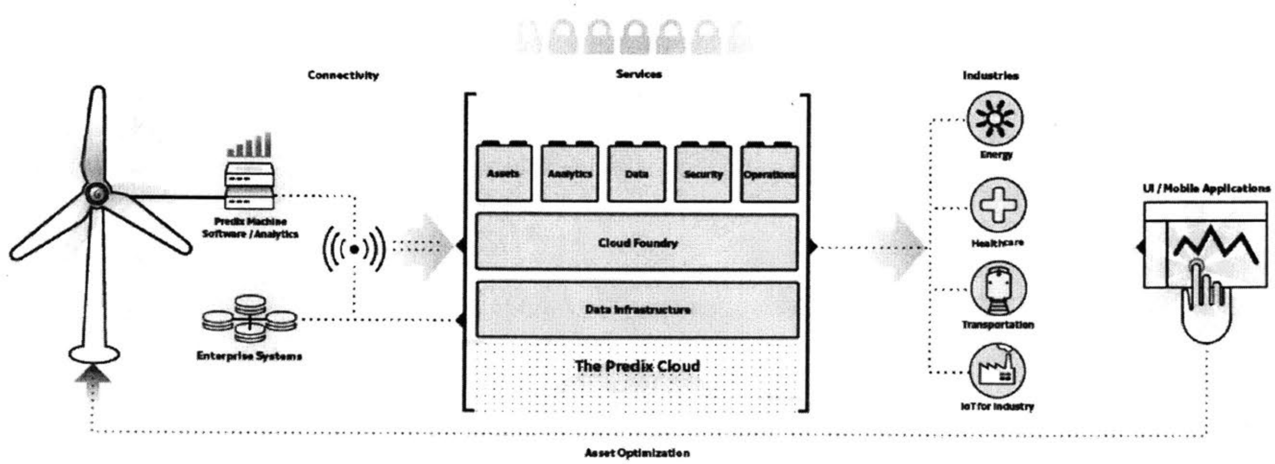

Figure 22. Architecture of Predix Platform (Source: [37])

To get hands-on experience with this cloud-based platform, I learned through practice on Predix. To gain full access to Predix, I created a free trial account. To launch a Predix-enabled development environment to streamline my development effort, I got familiar with the different development tools needed for different applications. To build an operation environment for Predix, I downloaded and used tools as follows:

- Cloud Foundry CLI
- Git
- Java SE Development Kit (JDK)
- Node.js
- NPM
- Maven
- Eclipse
- The Spring Tool Suite (STS) for Eclipse

To save time of installation, later, I used VirtualBox to run DevBox. DevBox is designed and tested with default settings to run up to four separate instances on my machine concurrently, depending on the capacity of the system. I have full root access, so I can extend and reconfigure the system. Through practice, I found Predix was an open cloud platform. It was easy to contribute to this ecosystem as an individual developer.

Industrial automation competitors have built their cloud-based platforms. However, not all platforms are open system. The comparison among industrial platforms is shown in the table 4 below.

Table 4. Primary Components of Predix

	Predix	PlantWeb	FactoryTalk
Company	GE	Emerson	Rockwell
Network Area	Internet (Cloud)	Intranet	Intranet
Attribute of Function	Ecosystem	Platform	Platform
Attribute of System	Convergence of OT and IT System	OT System (DCS/PLC/SCADA)	OT System (DCS/PLC/SCADA)
Development Openness	GE and Non-GE Companies (Startups)	Emerson only	Rockwell only
Provide security service for the public (developer)	Yes	No	No

2.2.7 Edge Computing and Device

Edge computing is a method of optimizing cloud computing systems "by taking the control of computing applications, data, and services away from some central nodes (the "core") to the other logical extreme (the "edge") of the Internet" which makes contact with the physical world. In this architecture, data comes in from the physical world via various sensors, and actions are taken to change physical state via various forms of output and actuators; by performing analytics and knowledge generation at the edge, communications bandwidth between systems under control and the central data center is reduced. Edge Computing takes advantage of proximity to the physical items of interest also exploiting relationships those items may have to each other. Many principles of Physics exhibit locality whereby an effect is greatest nearby and diminishes with distance. Edge Computing is the only form of Cloud computing that can offer "Proximity as a Service".

--Wikipedia [38]

Although cloud computing dominants the digital industry, edge computing is catching our eyes as in the digital transformation. Based on the location for data analytics, edge is divided into three classes, as is shown in the table 5 below.

Table 5. Classification of Edge

Class	**Location for Data Analytics**
Extreme edge	Sensor
Very edge	Machine/Equipment
Fog	Infrastructure

In the digital industry, edge locations are more likely refers to sensors, machines, equipment, ICSs, i.e. extreme edge and very edge that away from the cloud. Edge computing is suitable for devices with time-sensitive data and low latency requirement. Edge computing is likely to execute in small footprint devices such as a sensor hub or a gateway. Data analytics and machine learning algorithms are executed at edge. Please see the figure 23 below for the architecture of computing from edge to cloud.

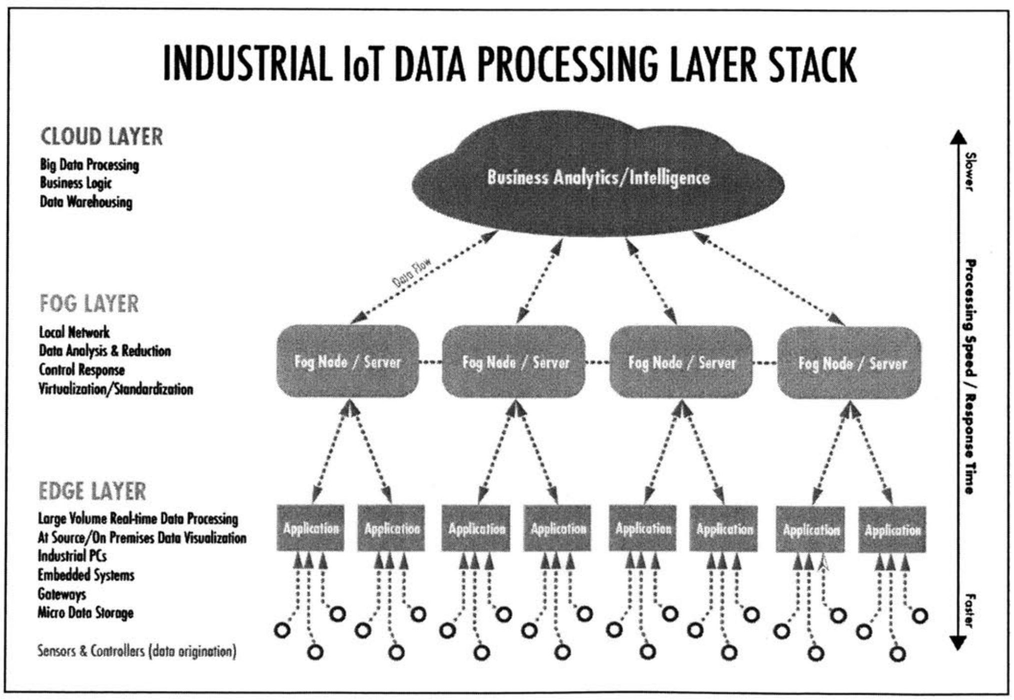

Figure 23. Architecture of Computing from Edge to Cloud (Source: [39])

Actually, to some extent, edge computing is similar to decentralized or distributed computing, which is not new to us. Digital industrial use cases give this technology new meanings. Analyzing huge amounts of machine-based data near the place where data come from in a scenario with a need of low latency is edge computing's advantage in the digital industry scenario.

Edge computing has potential value in use cases with low and intermittent connectivity. The use case characteristics are shown in the figure 24 below.

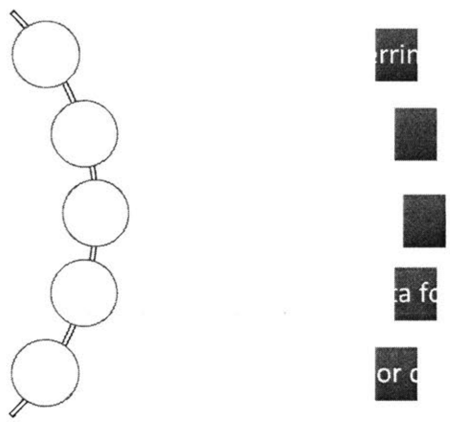

Figure 24. Characteristics of Edge Computing Use Case (Source: [40])

Some argue that edge computing will replace cloud computing in the future. However, to better create value of the huge amounts of data from machines and devices, edge computing and cloud computing should work together. Please see the figure 25 below for the cooperation between edge and cloud.

Edge computing is likely to take a more dominant position in scenarios with a need for low latency or with constraints in bandwidth, while cloud computing will create more value when using significant computing power or managing data volumes from across factories.

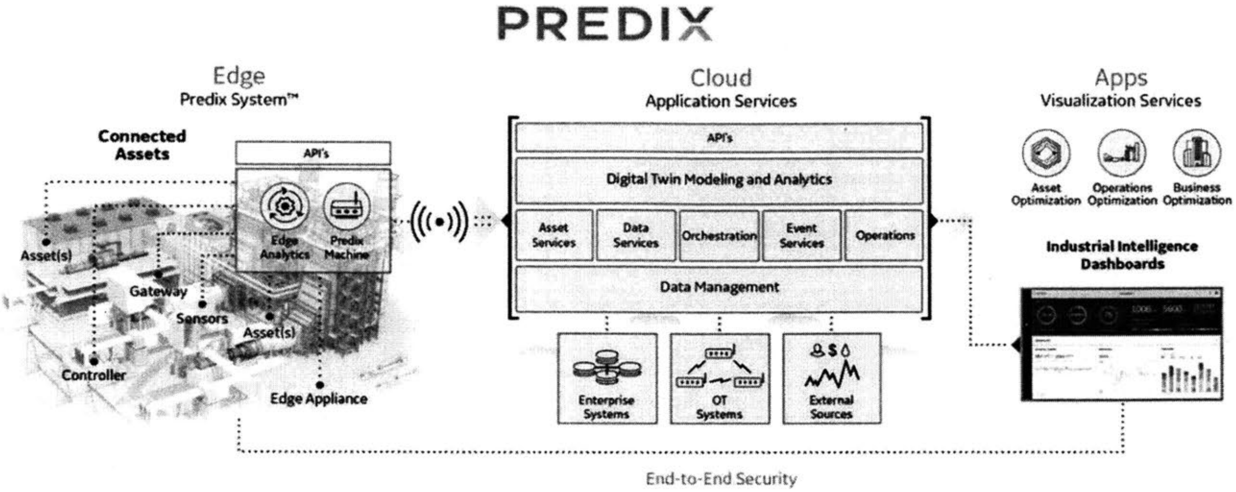

Figure 25. Cooperation between Edge and Cloud (Source: [40])

2.2.8 Big Data and Analytics

Many industrial companies and consulting companies predicated the number of device that will be connected in the future. Please see the table 6 below for the predication results.

Table 6. Predication of number of device that will be connected

	Billion devices by 2015	Billion devices by 2020	Billion devices by 2025
Gartner	4.9	20.8	
IDC			80
ABI		41	
Cisco			50
IBM	1000		
Intel		31	
Morgan Stanley		75	

Simply put, huge amount of data will be analyzed in the not-too-distant future. Industrial data are included. A large-scale factory produces over billion amounts of data. However, many raw data have no meaning, and lead to high latency and bandwidth.

Based on IBM's research, big data stand out in four dimensions: volume, variety, velocity and veracity (Four V's), as is shown in the figure 26 below.

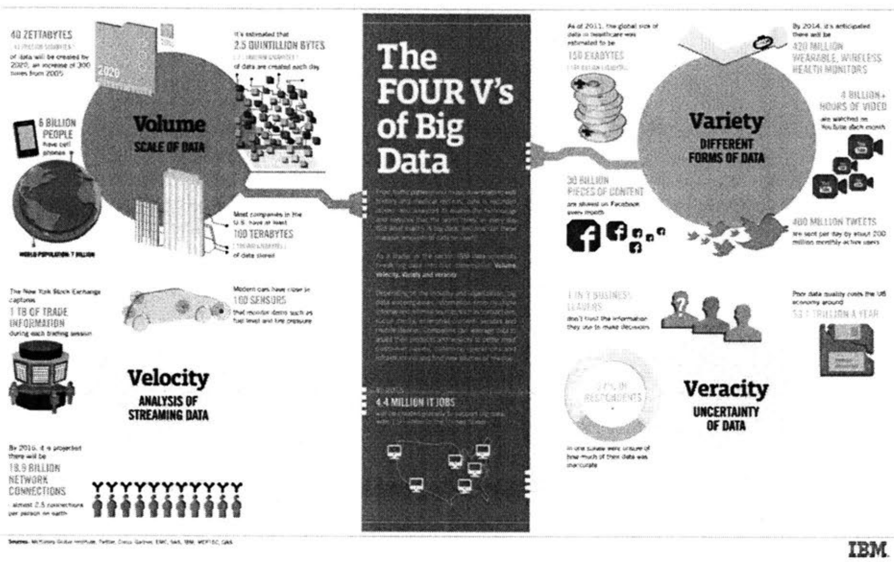

Figure 26. Four V's of Big Data (Source: [41])

Compared with internet data, industrial data are mostly made by machines with finer granularity and more complicate architecture. We value industrial data more on their comprehensiveness, mixture, relevancy, accuracy, low error-tolerant rate, and more importantly, the physical significance of features. For industrial data, clarifying requirements and logics and transforming them to mathematic models is essential. Therefore, data processing is more important than data analytics in the digital industry.

Besides relational database (RDB) and distributed database (DDB), time-series database (TSDB), which is also referred as to real-time database (RTDB), is more suitable for industrial data because most industrial data are structural time-series data, and these data can be searched according to time, date, and region. In addition, TSDB is more powerful because of its advantages in real-time read, write, and storage capability for huge amount of data and data collection capability with integrated industrial interfaces and protocols.

2.2.9 Artificial Intelligence (AI) and Machine Learning

AI is the new electricity.

--Andrew Ng, one of the most influential AI scientist in the world, founder of Google Brain

Various AI technologies such as computer vision, nature language processing (NLP), reasoning and optimization are changing our world over time, which is illustrated in the figure 27 below.

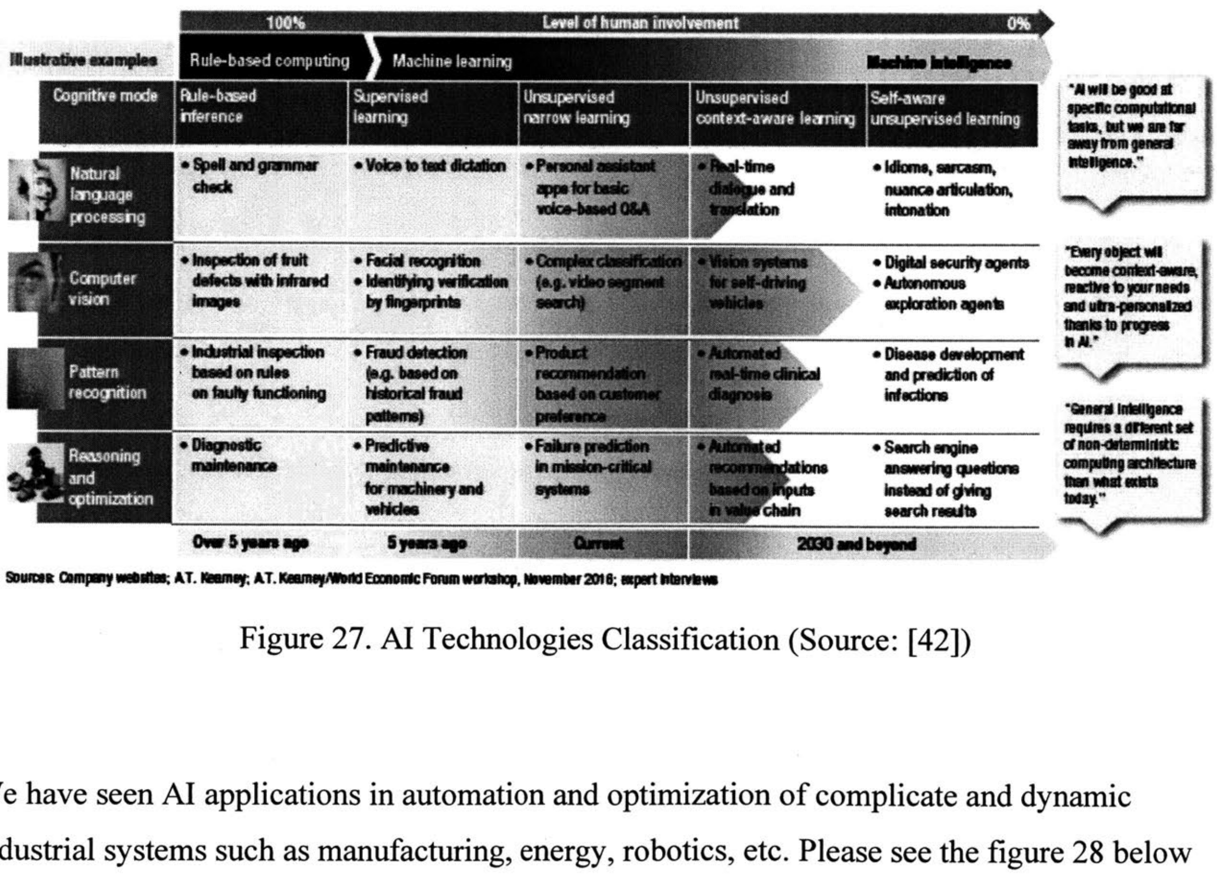

Figure 27. AI Technologies Classification (Source: [42])

We have seen AI applications in automation and optimization of complicate and dynamic industrial systems such as manufacturing, energy, robotics, etc. Please see the figure 28 below for the use cases with requirements and challenges.

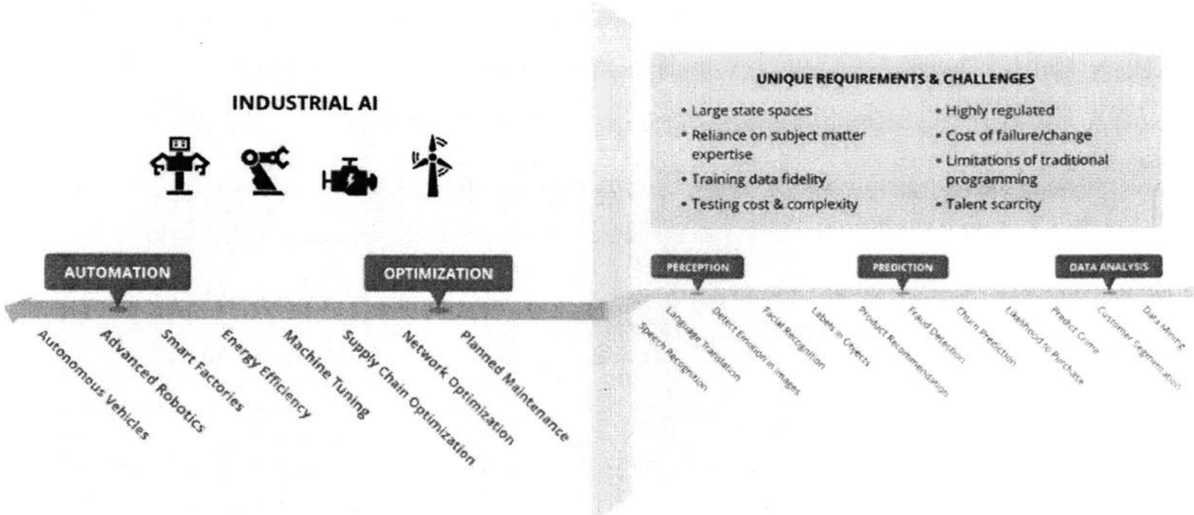

Figure 28. Industrial AI with Requirements and Challenges (Source: [43])

Without fundamental technologies for protocol interpretation, data collection, standardization, processing, or analysis from edge to cloud, AI can do nothing. Without verticals knowhow, industrial scenario understanding, or ability to translate production requirement, AI can do nothing. Without basic industrial automation hardware or software, autonomous (AI) can hardly bring its potential into full play.

What industrial customers need are end-to-end solutions rather than single products. Solutions are for equipment level or factory level. AI application in equipment level is the basis of that in factory level. It is essential to find out appropriate solutions for appropriate use cases, especially in equipment level at the beginning of the industrial AI age.

Also claimed by Andrew Ng, AI is changing the industry through adaptive manufacturing, automated quality control, and predictive maintenance. I introduce some use cases.

Quality control through computer visual inspection in manufacturing is a basic use case of AI and machine learning. With deep learning algorithms, AI-enhanced computers are able to detect every single tiny dot defect on circuit boards or chips and differentiate them from small particles and scratches on camera lens, which is beyond the limits of human eye. In addition, AI is able to inspect objects with high speed and accuracy and keep doing the job without taking a break.

Predictive maintenance is a more valuable use case. With machine learning methods, computers are able to analyze equipment's sensor data, including temperature and pressure, and predict the remaining useful life of each piece of equipment in order to reasonably schedule planned maintenance, reduce unexpected breakdowns and cost, and raise productivity [44]. I take a deep dive into predictive maintenance in the chapters below.

Various machine learning algorithms are applied in the digital industry, to name a few, Logistic Regression, Random Frost, Support Vector Machine, Decision Tree, K-Neighbors, Gaussian Naive Bayes, etc. Please see the figure 29 below for the word cloud. With the rapid development of deep learning, multi-layer neural network optimizes models and achieves better results.

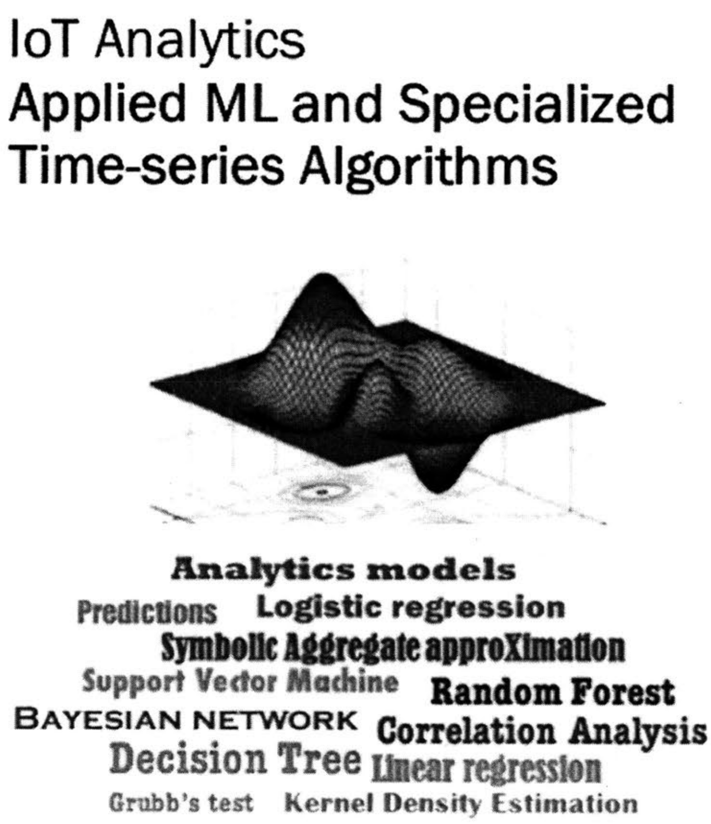

Figure 29. Word Cloud of Industrial AI Algorithms (Source: [45])

2.2.10 Other Technologies

Various digital technologies are emerging these years. Many of them such as AR/VR, Digital Twin, Cyber security, Block Chain, have potential value in the digital industry. Since these technologies are less associated with the predictive maintenance solution, I am not introducing them here.

2.3 Stakeholder Analysis

Based on the IIC reference architecture discussed above in the 2.1.2, stakeholders are divided into three classes: edge tier, cloud platform tier, and enterprise tier. I cite the figure 5 again below for convenience. Stakeholders can also be categorized into five domains: control domain, information domain, operation domain, application domain, and business domain. Flows are as shown in the table 7 below.

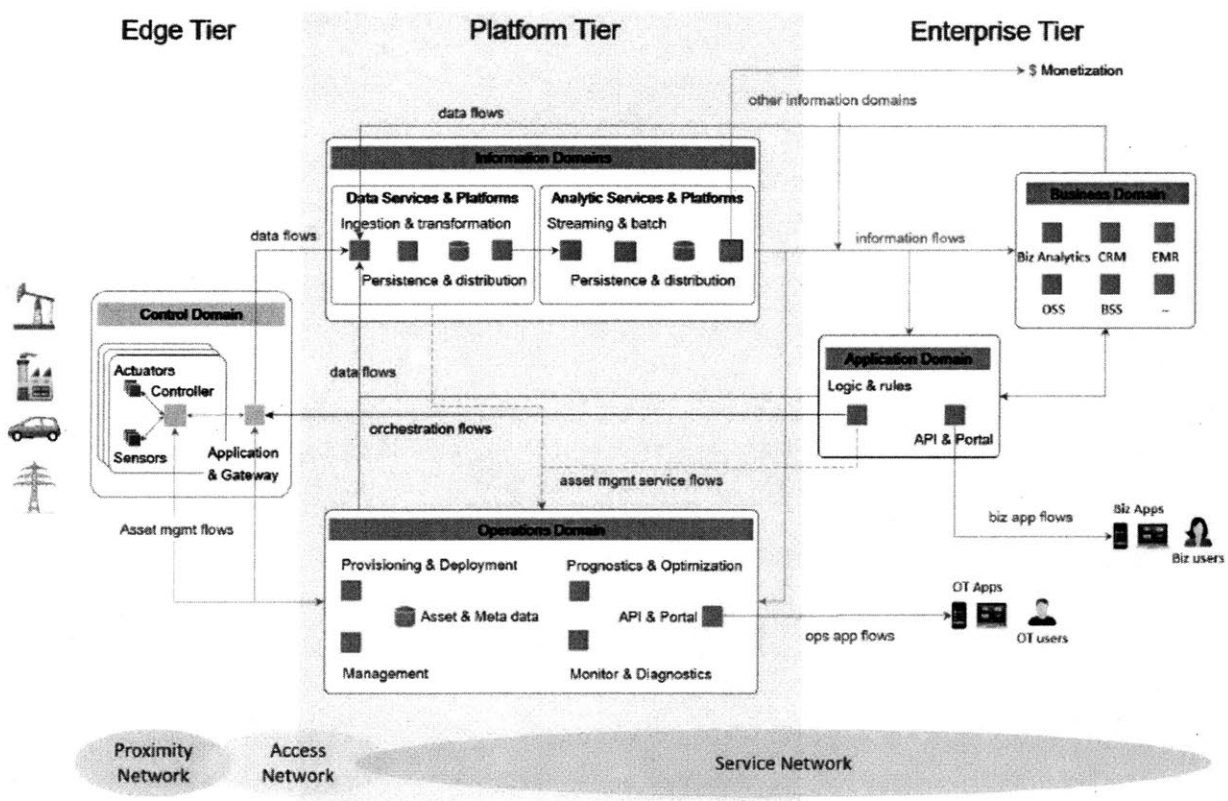

Figure 5. Mapping between a Three-tier Architecture to the Functional Domains (Source: [7])

Table 7. Relations among flows and domains

Flows/Domains	Control	Information	Operation	Application	Business
Asset mgmt flows	X	X	X		
Data flows	X	X	X		
Orchestration Flows	X			X	
Information flows		X	X	X	X
App flows				X	X

Stakeholders can also be categorized into five domains based on technologies they offer. Please see the figure 30 below.

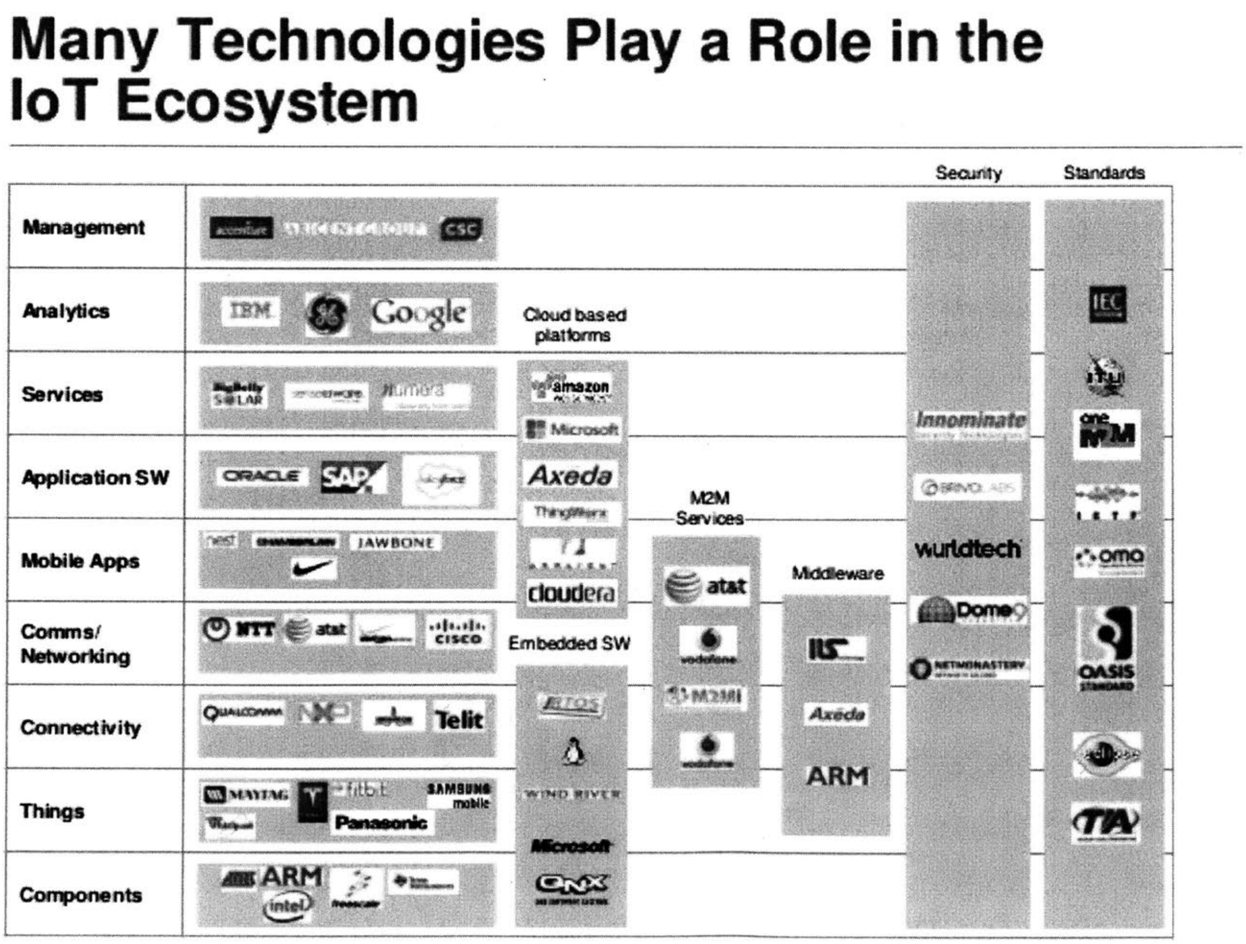

Figure 30. Stakeholder Classification based on Technology (Source: [46])

There are also stakeholders in regulatory side. Industrial companies must comply with the various applicable federal and state regulations. It is better to take into account the current and potential regulatory impacts. These agencies will demand compliance, environmental protection and safe operations.

2.4 Market Analysis

2.4.1 Market Trend

Based on data from MarketsandMarkets, the Industrial IoT market was valued at USD 113.71 Billion in 2015 and is estimated to reach USD 195.47 Billion by 2022. Based on a research by IndustryARC, the IIoT market will reach 123.89 billion USD by 2021. The revenue estimation is shown in the figure 31 below. Based on data from i-scoop, the IIoT market size and impact estimation is shown in the figure 32 below.

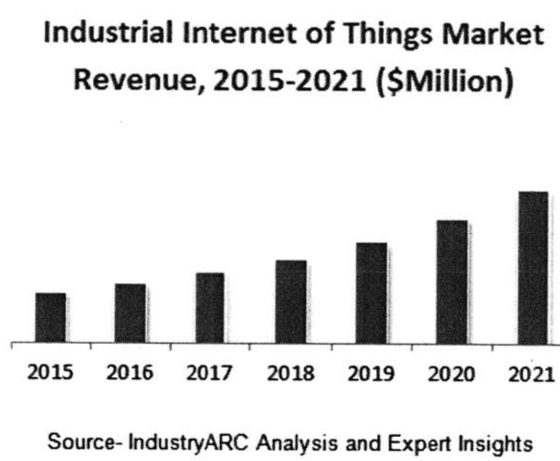

Figure 31. IIoT revenue trend (Source: [47])

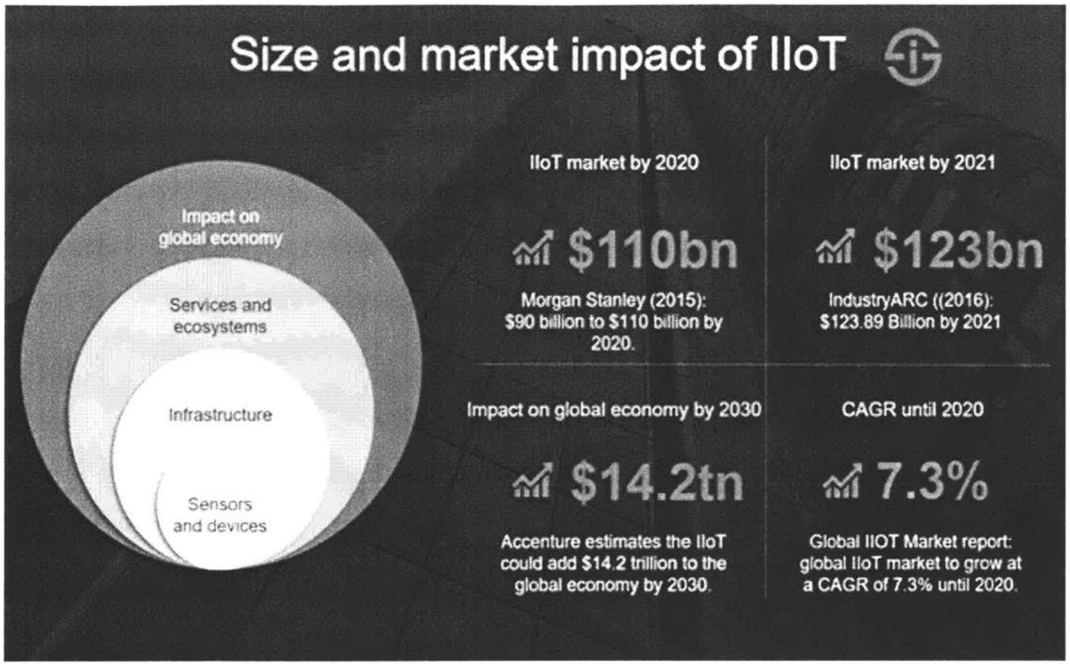

Figure 32. IIoT Market Size and Impact Trend (Source: [48])

2.4.2 Market Analysis by Vertical

Based on a search by MarketsandMarkets, the industry 4.0 market is divided into verticals below [49]:

- Automotive
- Aerospace
- Industrial Equipment
- Electrical & Electronics Equipment
- Chemicals & Materials
- Food & Agriculture
- Oil & Gas
- Energy & Power
- Healthcare
- Others (Pharmaceutical; Metal & Mining; Paper, Pulp, & Packaging; Water & Waste Water; Foundry & Forging; Textile & Cloth; Precision & Optics)

Based on a survey by i-scoop, manufacturing, including industrial equipment, electrical & electronics equipment, is the # 1 industry for IIoT in 2016, and seems to keep growing fast in the not-too-distant future. Please see the figure 33 below for the details.

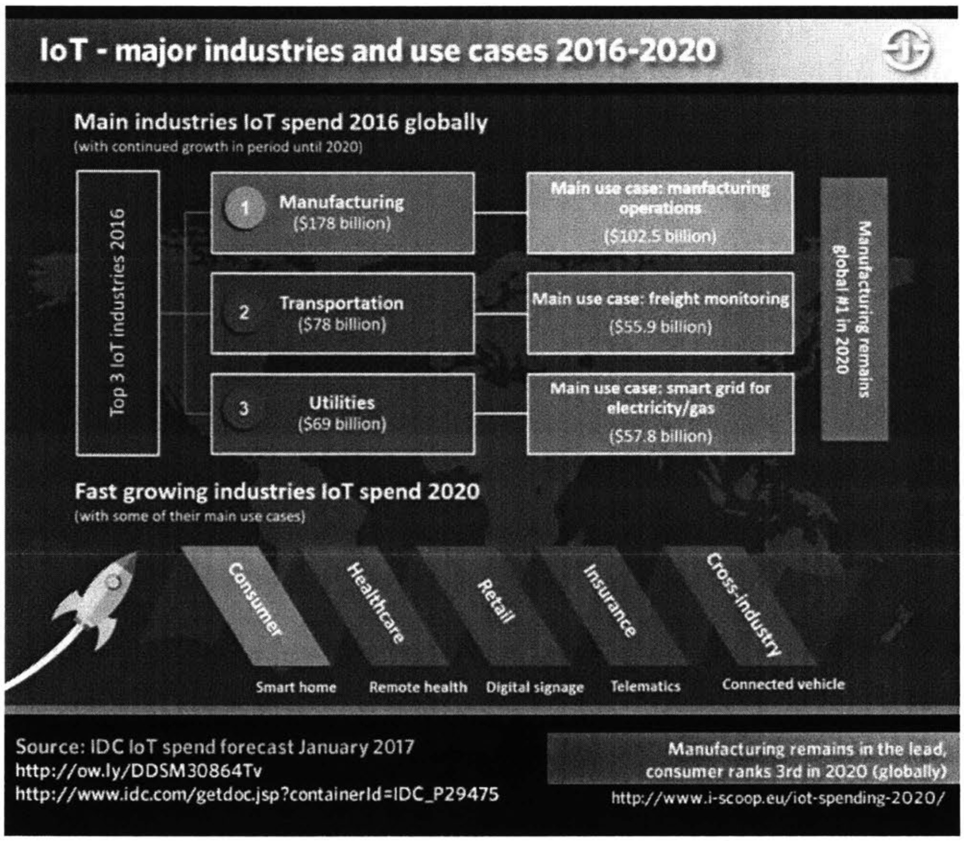

Figure 33. Main Industries in IIoT Market (Source: [50])

2.4.3 Market Analysis by Product/Service

Based on product or service, the digital industry market is divided into many sectors, such as sensors, edge devices, platforms, networks, analytics application, intelligence application, security application, digital twin, etc. Most major industrial companies have made investment on platform, application and digital twin. Please see the figure 34 below. Many startups work on stacks in specific verticals to survive in the competition with major companies. Please see the figure 35 below.

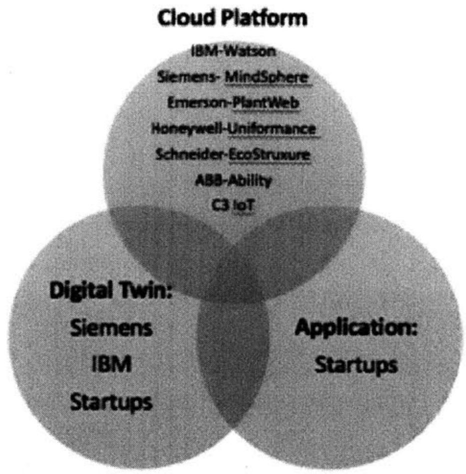

Figure 34. Major companies focused sectors

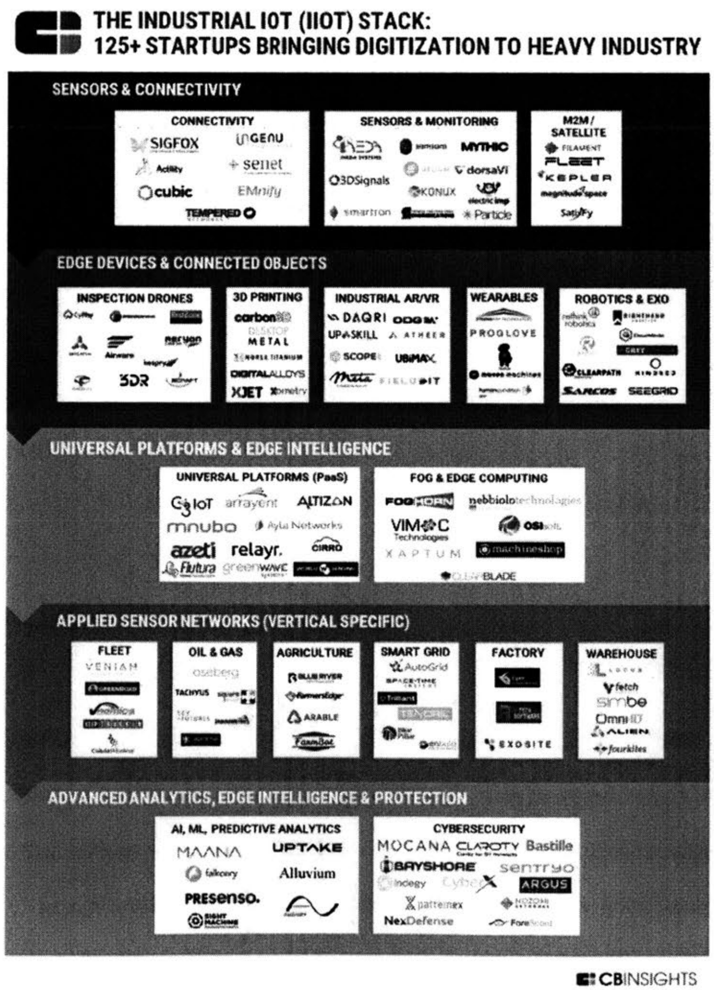

Figure 35. Startup Focused IIoT Stacks (Source: [51])

2.4.4 Market Analysis by Geography

Based on data from MarketsandMarkets, the commercialization of the IIoT applications is expected to be in the introductory stages in the developing countries in the APAC region. Countries such as China, Japan, South Korea, and India are taking initiatives to motivate the IIoT implementation. The dense population and the growing per capita income of the APAC region as well as large-scale industrialization and urbanization are dramatically driving the growth of the IIoT market. Please see the figure 36 below for the IIoT market by region. Major IIoT players are shown by geography in the figure 37 below.

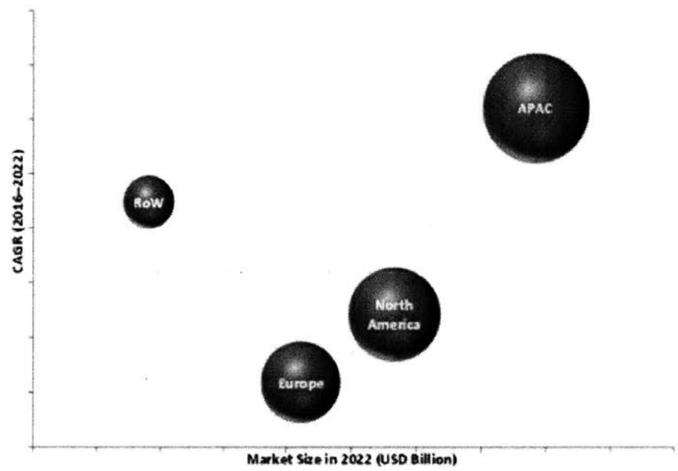

Figure 36. Industrial IoT Market by Region, 2022 (Source: [52])

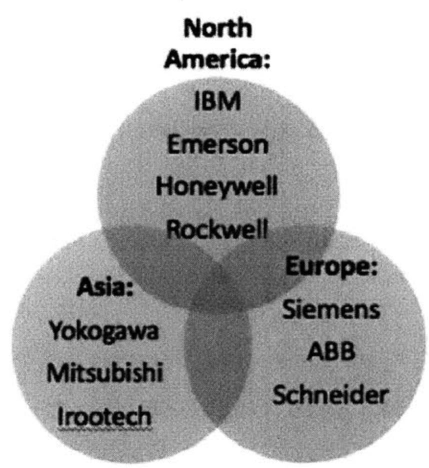

Figure 37. Major IIoT Players by Geography

2.5 Strategy Analysis

2.5.1 Product vs Service

Product or service, that is a question having been discussed for a long time. For hardware vendors, physical products are the most essential and valuable offerings, on the contrary, software providers value more on service. Software product and service analysis in various dimensions is shown in the table 8 below.

Table 8. Product vs Service (Source: [53])

Product Vs Service

	Software Product	Software Service
Cost of making	Design & Development the biggest (~80%) of the cost for software products. Hardware products have raw materials, processing and assembly costs too.	Salary, and minimal training costs. Resource is billed on time spent in a project. Can be in one project at a time.
Labor	Few but talented people required. Cost is high, returns are much higher.	Large number of people for more effort. Not everyone need to be highly skilled. Keeping cost low is a goal! Managing "bench" is raised to an art!
Revenue	Sell packaged software with license to use; Replicate into as many packages as required and keep selling. Revenue is independent of the cost of making the product.	Revenue is linearly proportional to number of people billed. To increase revenue - Increase billing rate; or add more people – "invent work"!
Force multiplier	A software product can be sold to millions by cloning, keeping the per-user cost affordably small.	Can't deploy a resource in multiple projects simultaneously. We don't have clones!
Cost reduction for customer	Not a concern. Replication cost is negligible anyways.	Hire cheaper talent, reduce timelines, reduce quality of deliverables. It hurts!
Innovation	Innovation is the essence of product development.	Not always required.

Copyright (c) Rajesh Varma, 2013 12

In the digital industry, there is a growing trend in Software-as-a-service. As an industrial equipment provider, GE announced to change itself to be a software company, and made significant investment on the digital transformation. Competitors such as Siemens, Honeywell, Schneider are becoming digital-driven. Therefore, a product-service hybrid offering is widely accepted in the digital industry.

2.5.2 Platform vs Application (Horizontal vs Vertical)

GE promoted Predix as Platform-as-a-service (PaaS) model in order to occupy the IIoT market in a horizontal way with the first mover advantage. Also, competitors mentioned above gradually took similar actions.

To illustrate it, GE (Digital) and Siemens (Digital Factory) built their cloud-based IIoT platform, and provided applications for their wind turbines, aircraft engines, and medical equipment in vertical industries; Many industrial automation companies, such Emerson, Honeywell, Schneider built their IIoT platform with support from cloud provider especially Microsoft. IBM (Watson) and ABB (Ability) not only built platforms but also teamed up for Industrial AI; Through acquisition and building itself, Cisco owns three IIoT platforms: Jasper, Kinetic, and a hardware platform. Please see the figure 38 below for the major IIoT platform providers by verticals.

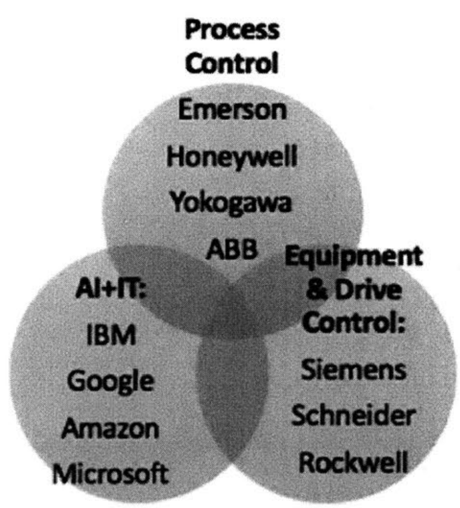

Figure 38. Major IIoT Platform Provider by Verticals

Additionally, many other players built platforms to secure their markets in their industries, such as IT, Internet, Telecom, etc. Platforms are emerging in various models, such as Infrastructure

Cloud Platform (ICP), Connection Management Platform (CMP), Device Management Platform (DMP), Application Enable Platform (AEP), Business Analytics Platform (BAP), etc.

There are many opportunities for startups that provide platform, application and service in IIoT. C3 IoT and Uptake are providing predicative platform and bespoke consulting service to industrial customers. Foghorn system focuses on edge intelligent platform and applications. Collaborating with Tencent, SANY's subsidiary – Irootech is developing the cloud platform in China. Many more startups are looking for specific verticals and regional markets and pilot project opportunities. Being acquired by major companies might be a good approach for many startups.

This year, Gartner defined the market for Industrial Internet platforms as a set of integrated software capabilities and provided evaluation: Magic Quadrant for Industrial IoT Platforms. Please see the figure 39 below. The evaluation criteria are shown in the figure 40 below.

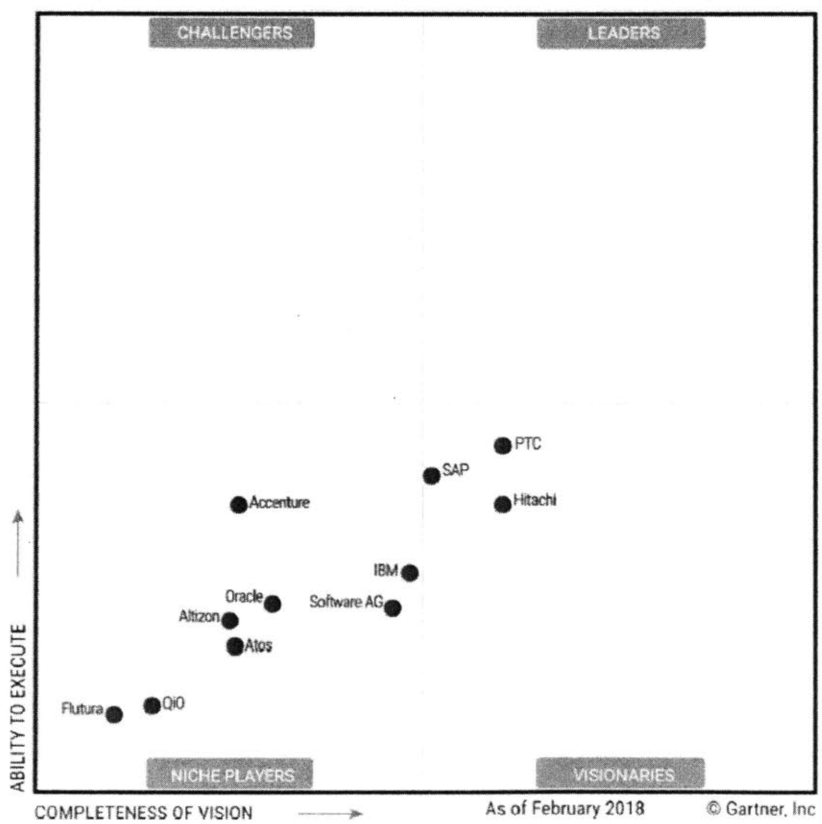

Figure 39. Magic Quadrant for Industrial IoT Platforms (Source: [54])

Figure 40. Magic Quadrant for Industrial IoT Platforms (Source: [54])

However, significant investment on platform might not be a reasonable trend. Many platforms are up in the air. A lack of essential domain knowledge is a big challenge for these platforms. What's worse, due to a lack of organized data, practical vertical experience, and appropriate policy, many pilot projects have not created sufficient value as expected, many companies have encountered difficulties in the adoption and implementation of digital technology. GE sold its industrial solution business unit to ABB. Predix is facing similar challenge. Therefore, horizontal strategy is not suitable for most companies in this scenario at this time.

Through reflection, more stakeholders realized the importance of domain expertise and applications in specific verticals. Predictive maintenance applied to verticals is more valued. This is also a reason why I want to take a dive deep into this technology in the following chapters.

Chapter 3: Solution Analysis

Maintainability is defined as the probability that a system or system element can be repaired in a defined environment within a specified period of time. Increased maintainability implies shorter repair times.

--- American Society for Quality (ASQ) 2011

3.1 Background of Maintenance

In industry, most companies have to weigh lost production time against the risks of breakdowns. Based on a study by Emerson, unplanned downtime costs industrial manufacturers an estimated $50 billion annually [55]. A single pump failure can cost $100,000 to $300,000 a day in lost production. According to a survey by PTC, poor maintenance methods cost a factory's productive capacity between 5 and 20 percent [56].

There are four industrial maintenance methods: reactive maintenance, planned maintenance, proactive maintenance, and predictive maintenance. Comparison in concept among maintenance strategies is shown in the table 9 below. Their development and evolution over time are shown in figure 41 below. Obviously, predictive maintenance is the most efficient and promising solution. Based on data acquired from connected smart machines, time points and locations of failures might occur can be accurately and efficiently predicted, unnecessary downtime can be substantially minimized. Please see the figure 42 below for the comparison in operation among maintenance strategies.

Table 9. Comparison in Concept among Maintenance Strategies

Strategy	**Model**	**Description**
Reactive Maintenance	Run-to-failure	Only performing maintenance when problems occur
Preventive Maintenance	Regularly scheduled	Using either time intervals or usage as a trigger
Proactive Maintenance	Root cause analysis	Measures are taken to prevent equipment failure altogether

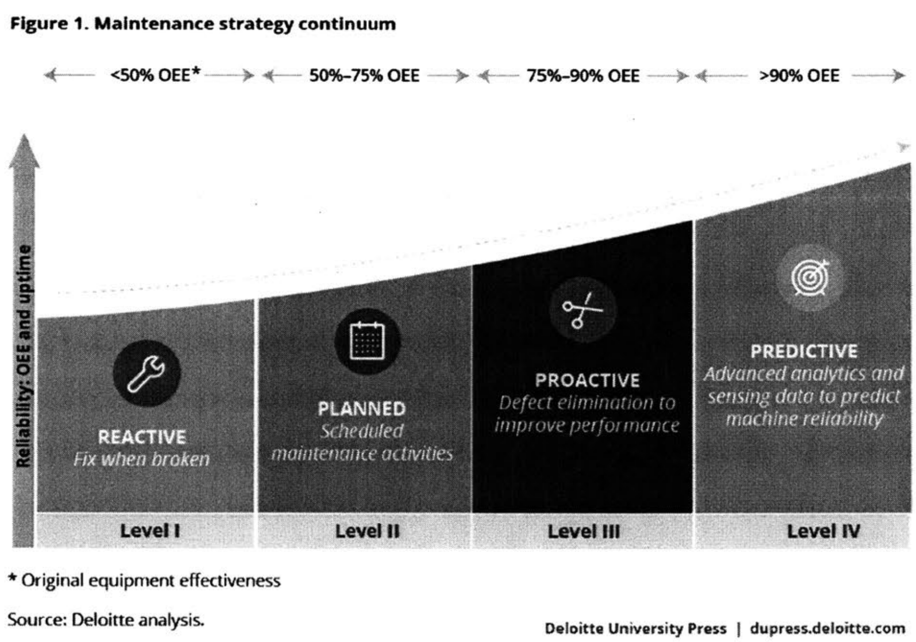

Figure 41. Maintenance Strategy Continuum (Source: [57])

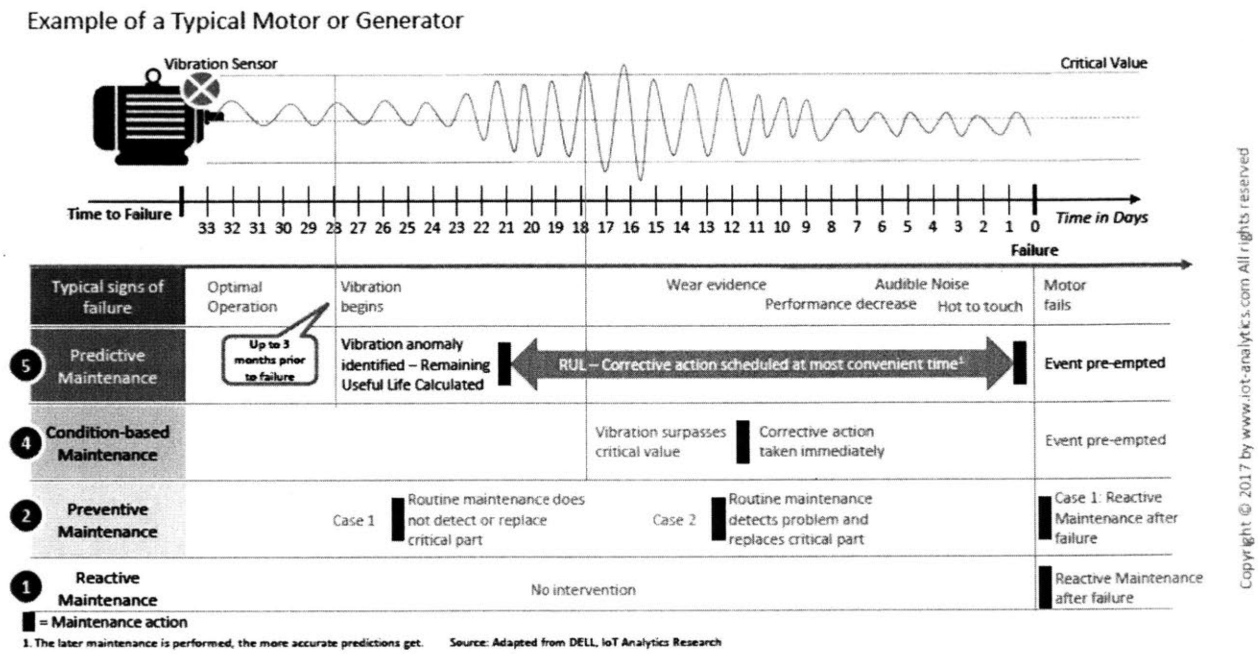

Figure 42. Comparison in Operation among Maintenance Strategies (Source: [58])

3.2 Introduction of Predictive Maintenance

Predictive Maintenance (PdM) is a practical technology that helps evaluate the equipment condition to predict time when the equipment fail to make preparation for maintenance. Predictive maintenance helps evaluate in-service equipment's condition to predict time when the equipment fail and make preparation for maintenance, avoiding unplanned breakdown and downtime. To reduce uncertainty and manage risk, this technology could be used in many scenarios in industries.

The importance of PdM and related analytics is rapidly growing in these years. Based on a report from IoT Analytics, PdM becomes the most important application in industrial analytics in the next 1-3 years. Please see the figure 43 below for the survey result of importance.

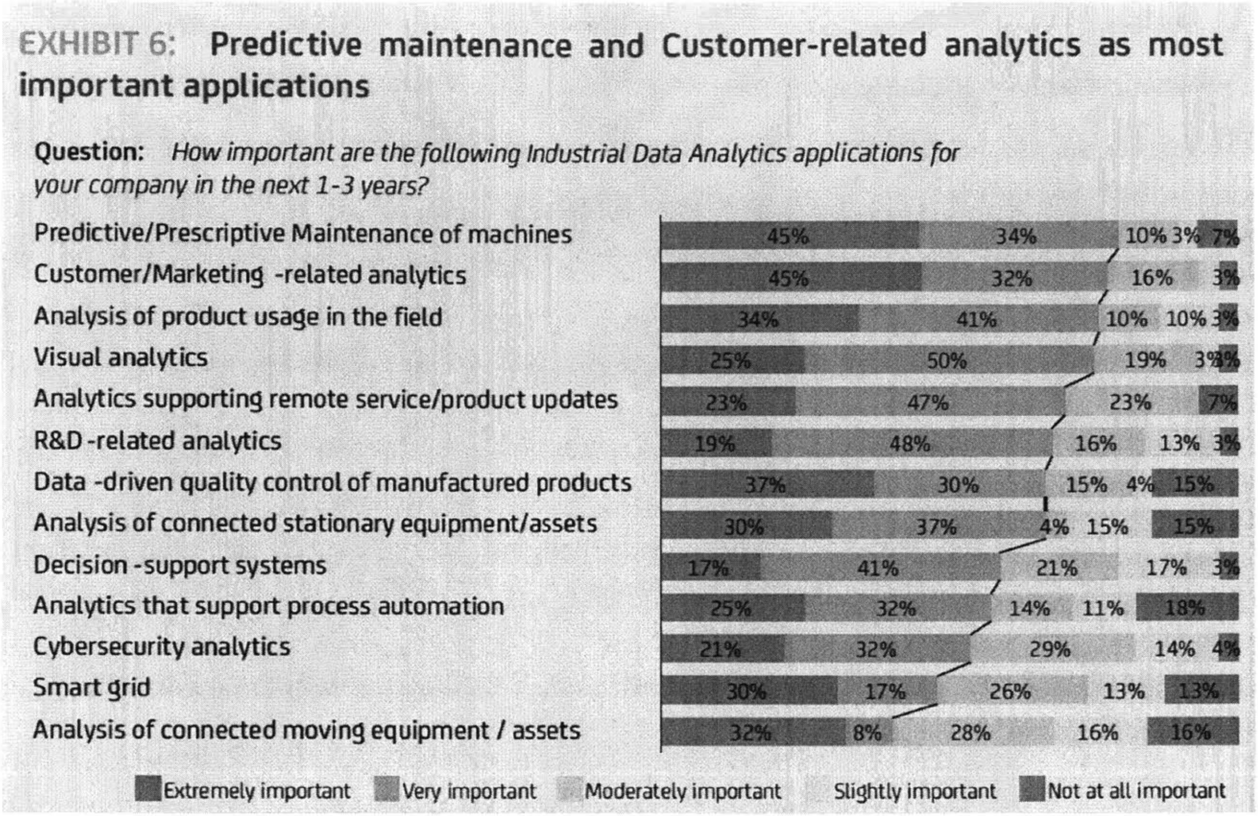

Figure 43. The Importance of PdM (Source: [59])

Suitable and unsuitable scenarios for PdM are shown in the table 10 below.

Table 10. Suitable and Unsuitable Scenarios for PdM (Source: [60])

Suitable scenarios	**Unsuitable scenarios**
Have a critical operational function	Do not serve a critical function
Have failure modes that can be cost-effectively predicted with regular monitoring	Do not have a failure mode that can be cost-effectively predicted

3.3 Market Analysis

IoT Analytics Research Team published Predictive Maintenance – Market Report 2017-2022, systematically reviewed a change from Condition-based Maintenance to IoT-& Analytics-Enabled Predictive Maintenance.

Based on this report, PdM use case ranks #1 in connected industry settings. PdM market revenue will reach $10.96 Billion by 2022. A compound annual growth rate is 39% in the rapid growing market. Maintenance efficiency achieve a 20%-25% increase in real project reports [58]. Please see the figure 44 below. The main driving forces of PdM market development are shown in the figure 45 below:

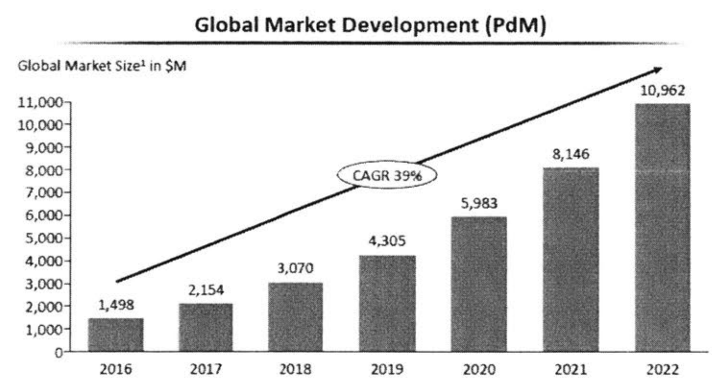

Figure 44. PdM Market Development (Source: [58])

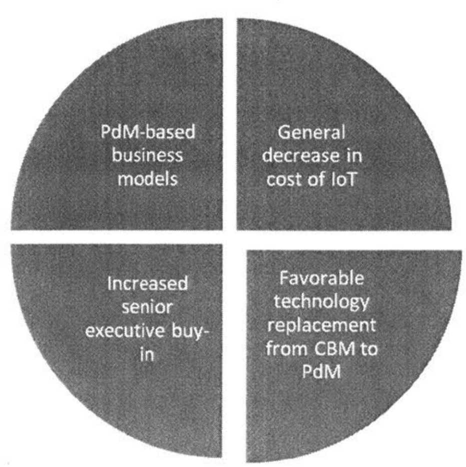

Figure 45. Driving Forces of PdM Market Development

3.4 Business Model Analysis

PdM helps manufactures reduces downtime and enables optimized maintenance events from operators and service providers. With more accurate and reasonable maintenance scheduling, PdM is able to maximize utilization of maintenance resources and optimize maintenance management regarding inventory and supply chains. PdM is able to minimize time the equipment is being maintained, the production hours lost to maintenance, and cost of spare parts, supplies, and human resource especially experienced personnel. IoT-enabled PdM is able to achieve over 20 percent of addressable costs as well as production and operational benefits. In addition, with PdM product and service, equipment providers strengthen their competitiveness in the digital industrial market [58].

3.5 Technology Review

Actually, PdM technology itself was originated from NASA. Later, PdM has been developed and promoted by industrial companies such as GE, Siemens, SAP, IT companies such as IBM, Microsoft, and companies in specific verticals. Please see the figure 46 below for the ranking of companies in PdM area.

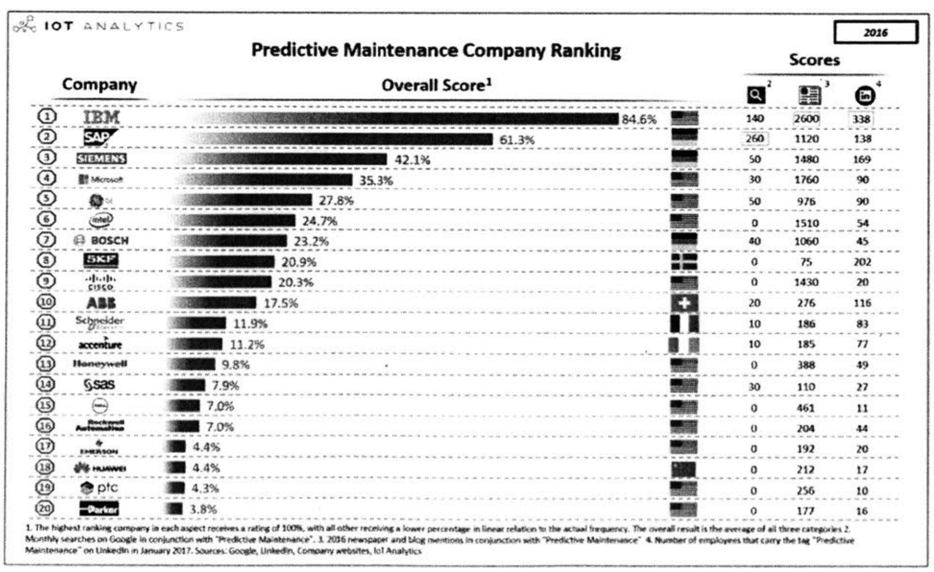

Figure 46. PdM Company Ranking (Source: [58])

To achieve PdM, technologies are required in six stacks: device, connectivity, storage, platform, analytics, application. Most of these technologies were introduced in the chapter 2. Hardware companies in PdM provide products and services with first three technologies, while software companies' offerings focus on last three technologies. In addition, AI companies' key technologies are in analytics and application sector. These PdM related technologies in the six stacks are considered at the Technology Readiness Levels (TRL) 8 or 9. Please see the definition of TRL by NASA and figure 47 below for the criteria of TRL.

Technology Readiness Levels (TRL) are a type of measurement system used to assess the maturity level of a particular technology. Each technology project is evaluated against the parameters for each technology level and is then assigned a TRL rating based on the projects progress. There are nine technology readiness levels. TRL 1 is the lowest and TRL 9 is the highest.

--NASA [61]

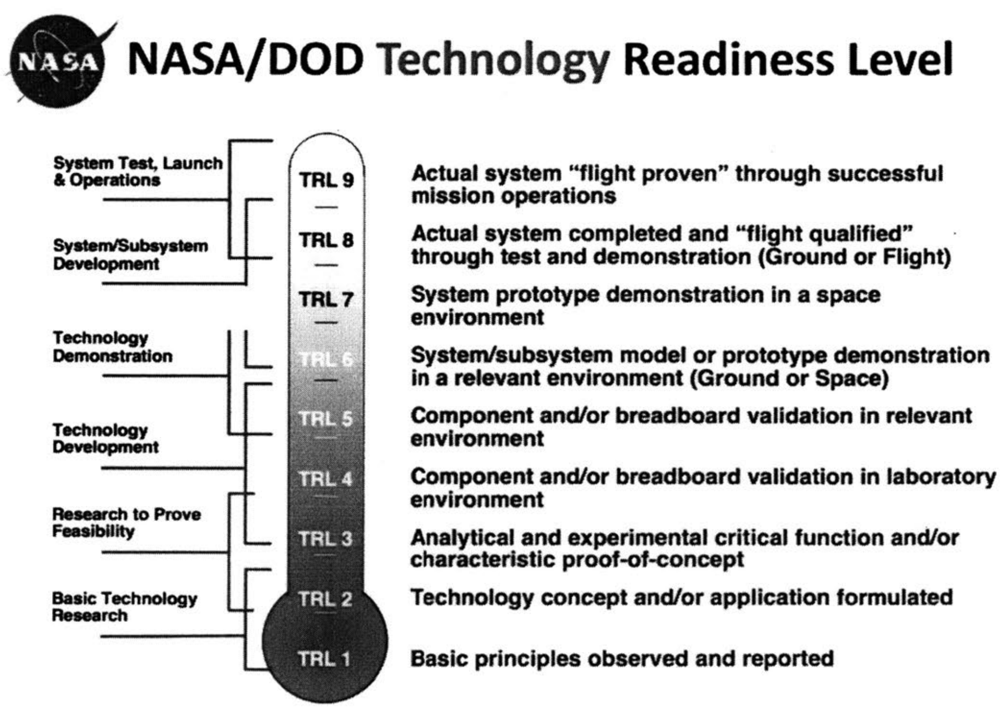

Figure 47. TRL Criteria (Source: [62])

Therefore, these technologies mature enough for application in industry. To better understand the predictive maintenance solution and obtain convincing results, I designed and implemented an experiment and obtained hands-on experience on machine learning.

Chapter 4: Experiment Design

4.1 Goal

I design an experiment to implement predictive maintenance with machine learning approach to evaluate the performance of failure prediction and ensure that this solution indeed optimize results and create value. Based on given data of sensors and completed cycles, I seek optimal model for engine degradation prediction: predict the number of remaining operational cycles, which is also referred as to Remaining Useful Life (RUL), prior to failure; and predict failure possibility of specific engine in next n-steps. Also, I would like to get hands-on experience of machine learning methodologies and tools through this experiment.

4.2 Data Source and Description

Getting industrial data is difficult because of their confidentiality. These datasets are published on the website of NASA's Prognostics Center of Excellence (PCoE) [63]. The data were originated from the Commercial Modular Aero-Propulsion System Simulations (C-MAPPS) system. The approach, solution, and datasets were used in the IEEE 2008 Prognostics and Health Management (PHM08) conference challenge problem [64] - [70].

The data are from a fleet of engines, each with its fault points, resulting in its degradation. These are multivariate time series data. Training and test subsets are included in each data set. All engines are of the same type. Each time series data is from each engine. Each engine has its initial wear and manufacturing variation that is not a fault condition. Three operational settings have impact on engine performance. Sensor noise is also existed in the datasets. Engines fail at certain time points during the series. Faults grow in magnitude until the system fails in the training set, while the time series ending happens before the system fails in the test set. Train trajectories, test trajectories, conditions, and fault modes in these datasets are as shown in the table 11 below. To further understand the data and the system, plcase see figure 48 for the simplified diagram of engine simulated and figure 49 for a layout of modules and connections [63].

Table 11. Statistical Data (Source: [63])

	FD001	FD002	FD003	FD004
Train trajectories	100	260	100	248
Test trajectories	100	259	100	249
Conditions	1	6	1	6
Fault modes	1	1	2	2

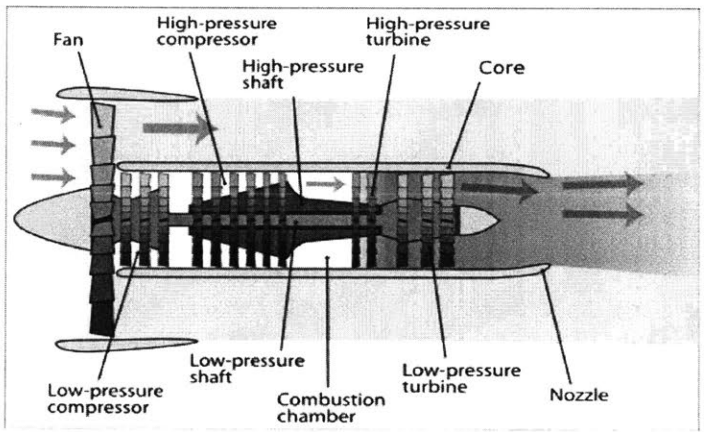

Figure 48. Simplified Diagram of Engine Simulated (Source: [71])

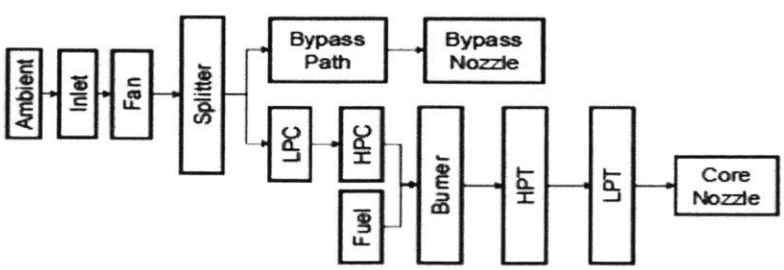

Figure 49. Modules and Connections Layout (Source: [68])

The data show initial wear and manufacturing variability of a set of engines. Each engine fails at some time point. The data also show operational settings of the turbofan in each cycle and a measurement of each of 21 sensors in that cycle. 26 columns of numbers are included in this data. Each row of data is taken during an operational cycle, while each column of data is a variable. Please see the table 12 below for sensors' physical meaning and the table 13 below for the name of each column in the raw data [63].

Table 12. Sensors' physical meaning (Source: [68])

Symbol	*Description*	*Units*
Parameters available to participants as sensor data		
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s
Parameters for calculating the Health Index		
T48 (EGT)	Total temperature at HPT outlet	°R
SmFan	Fan stall margin	--
SmLPC	LPC stall margin	--
SmHPC	HPC stall margin	--

Table 13. Name of each column in thc raw data (Source: [63])

1	Unit number
2	Time in cycles
3	Operational setting 1
4	Operational setting 2
5	Operational setting 3
6	Sensor measurement 1
7	Sensor measurement 2
8	……
9	Sensor measurement 21

4.3 Methods, Steps, and Tools

First, achieve the baseline approach, and make an evaluation of models based on performance metrics. The Steps in the baseline approach are shown in the table 14 below. Then, implement an advanced approach with deep learning methods. Finally, compare with results between baseline approach and optimal approach. Code would be written with Python on the Jupyter Notebook environment. Libraries and tools such as numpy, pandas, scikit learn, keras and tensorflow would be used.

Table 14. Steps in the Baseline Approach

Order	**Step**	**Description**
1	Data preparation	Use labeling, normalization and visualization to process and review data
2	RUL prediction	Use regression methods to predict RUL; Select and run appropriate models such as Linear, Decision Tree, Random Forest regression models; Evaluate models based on performance metrics such as Root Mean Square (RMSE).
3	Failure possibility prediction	Use classification method to predict failure possibility in next n-steps; Select appropriate models such as Decision Tree, Random Forest models; Evaluate models based on performance metrics such as Accuracy, Precision.

Chapter 5: Experiment Implementation

5.1. Baseline Approach

The baseline approach is fundamental to the entire experiment. The optimal approach is built on the baseline approach.

5.1.1 Data Preparation

Data preparation is the initial step before any data analysis. It includes data reading, data labeling, data visualization, and data normalization. The result could be used in both baseline approach and optimal approach.

5.1.1.1 Data Reading

Training data, test data, and ground truth data are loaded, columns with NaN data are dropped, and column names such as setting1, sensor1, sensor2 are added.

5.1.1.2 Data Labeling

The training data does not include RUL as a target variable. The only provided RUL is for the last cycle of each engine. The cycle numbers could be used to label training data with the equation below.

$RUL = Max\ Cycle - Current\ Cycle$

Columns of labels are added such as rul, w1. Given that, w1 steps in 30 remaining time series. See below for the Table 15 of labeled training data.

Table 15. labeled training data

	id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11
0	1	1	-0.0007	-0.0004	100.0	518.67	641.82	1589.70	1400.60	14.62	21.61	554.36	2388.06	9046.19	1.3	47.47
1	1	2	0.0019	-0.0003	100.0	518.67	642.15	1591.82	1403.14	14.62	21.61	553.75	2388.04	9044.07	1.3	47.49
2	1	3	-0.0043	0.0003	100.0	518.67	642.35	1587.99	1404.20	14.62	21.61	554.26	2388.08	9052.94	1.3	47.27
3	1	4	0.0007	0.0000	100.0	518.67	642.35	1582.79	1401.87	14.62	21.61	554.45	2388.11	9049.48	1.3	47.13
4	1	5	-0.0019	-0.0002	100.0	518.67	642.37	1582.85	1406.22	14.62	21.61	554.00	2388.06	9055.15	1.3	47.28

The RUL of test data depends on both the given max cycle in test data and real max cycle in ground truth data. The equation is shown below.

RUL=Given Max Cycle + Real Max Cycle – Current Cycle

The labeling process regarding w1 in test data is similar to that in training data.

5.1.1.3 Data Visualization

Please see the figure 50 below for one group of training sensor data over cycle time.

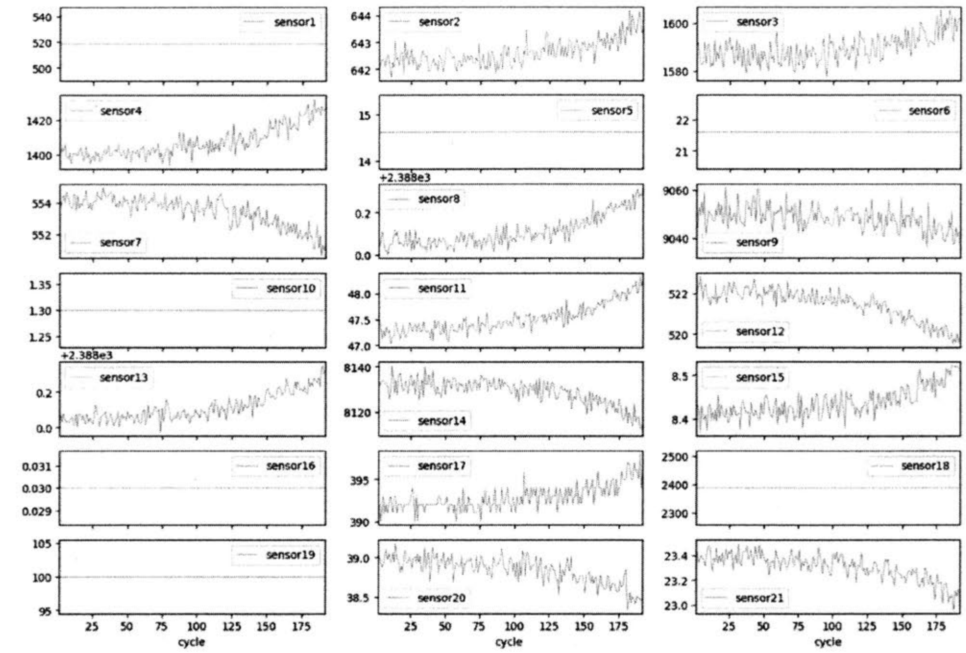

Figure 50. One Group of Training Data

Please see the figure 51 below for 10 groups of sensor data over training cycle.

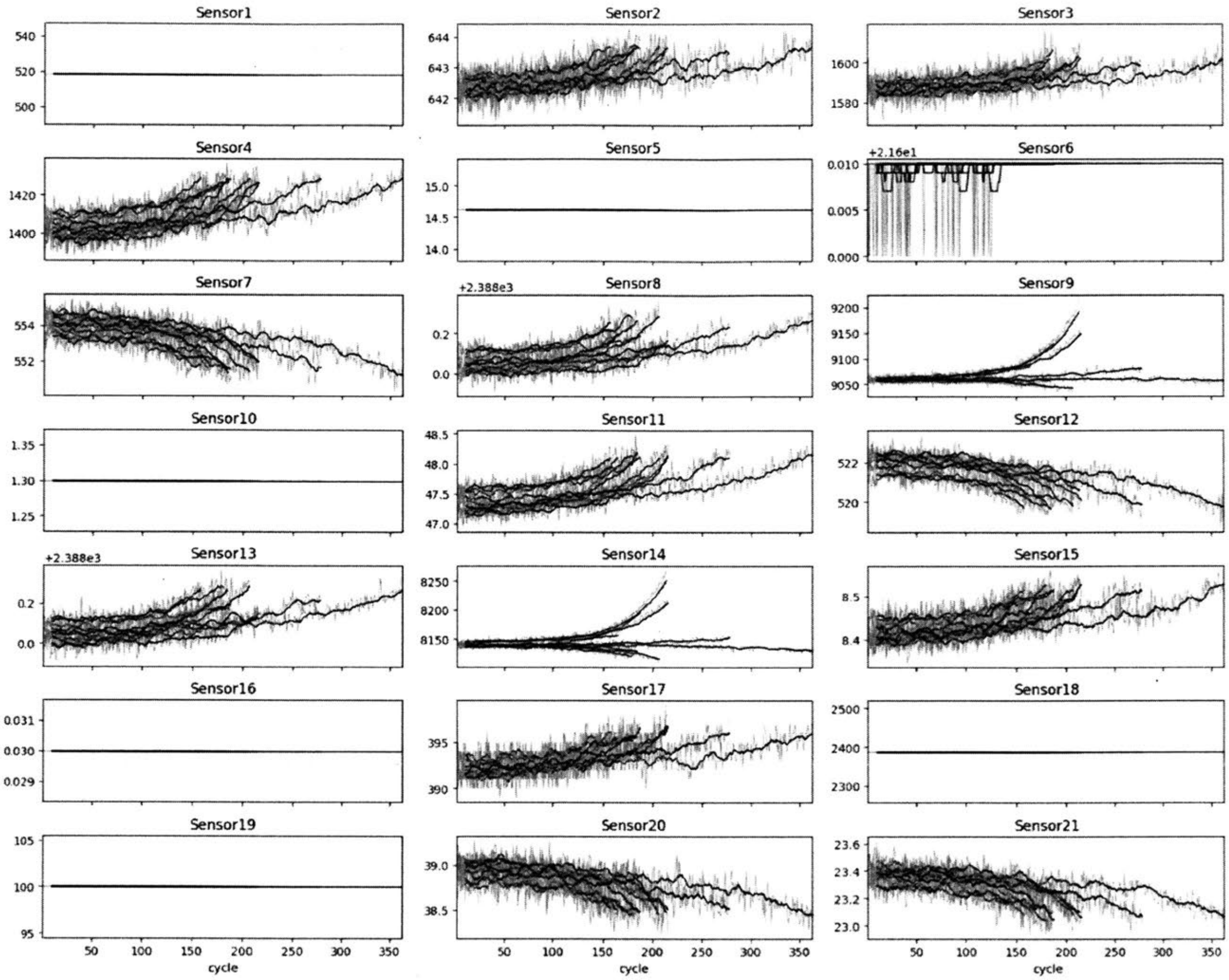

Figure 51. Ten Group of Training Data

10 groups of sensor data over RUL cycle is also observed. Sensor data fluctuates towards end of engine life. See the figure 52 below.

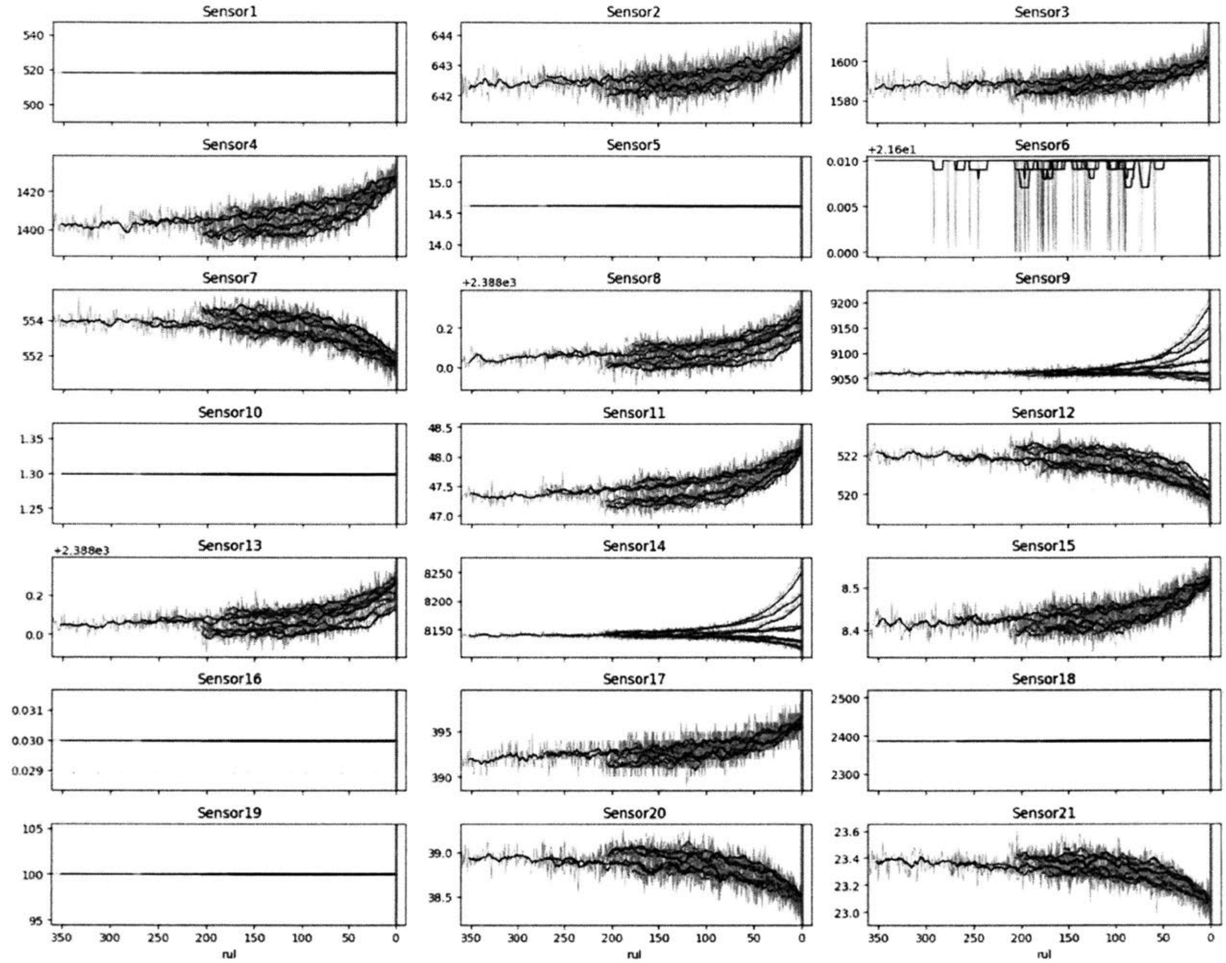

Figure 52. Ten Groups of Training Data over RUL

5.1.1.4 Data Normalization

Features are transforms by scaling each feature to a given range, such as between zero and one. MinMax normalization is used to linearly transform x to y= (x-min)/(max-min), where min and max are the minimum and maximum values in X, and X is the set of observed values of x [72].

5.1.2 Regression Model Selection and Evaluation

Regression models are as follows. Linear regression is a baseline model to predict RUL (target) from sensors measurements (predictors).

- Linear regression
- LASSO
- Ridge regression
- Decision tree regression
- Random forest regression

Regression performance metrics are as follows:

- Explained variance
- Root mean squared error
- Mean absolute error
- R2 score

One of the most important metrics for loss evaluation is the Root Mean Squared Error (RMSE), please see the equation below.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - yi^{\wedge})^{2}} \quad yi = predicted\ value, \qquad yi^{\wedge} = actual\ value$$

Please see the table 16 below for the comparison of regression models based on performance metrics.

Table 16. Comparison of Regression Model based on Performance metrics

	LinearRegr	Lasso	Ridge	DecisionTr	RandomFore
explained variance	0.337138	0.337516	0.337424	0.339755	0.351758
mean absolute error	37.163542	37.151975	37.154635	36.353867	36.766395
r2 score	0.327030	0.327385	0.327301	0.335055	0.330541
root mean squared error	48.382349	48.369585	48.372582	48.093018	48.255980

Please see the figure 53 below for the feature importance analysis. Based on feature importance scores, some sensors' data might be dropped to lower the loss.

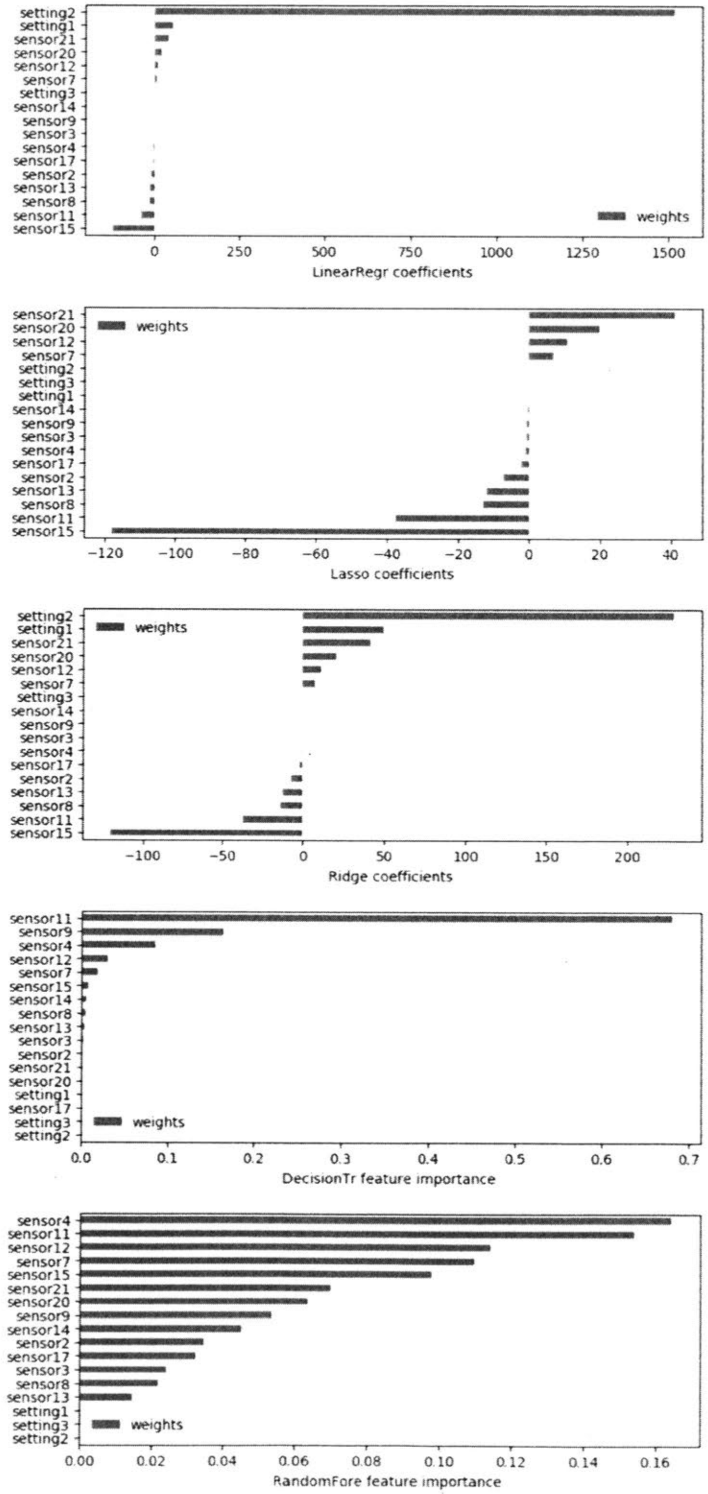

Figure 53. Feature Importance Analysis

Please see the figure 54 below for the feature coefficients weights. In the feature correlation heatmap, feature 9 and 14 are strongly correlated (0.95), and feature 9 only has strong correlation with 14, not with the rest.

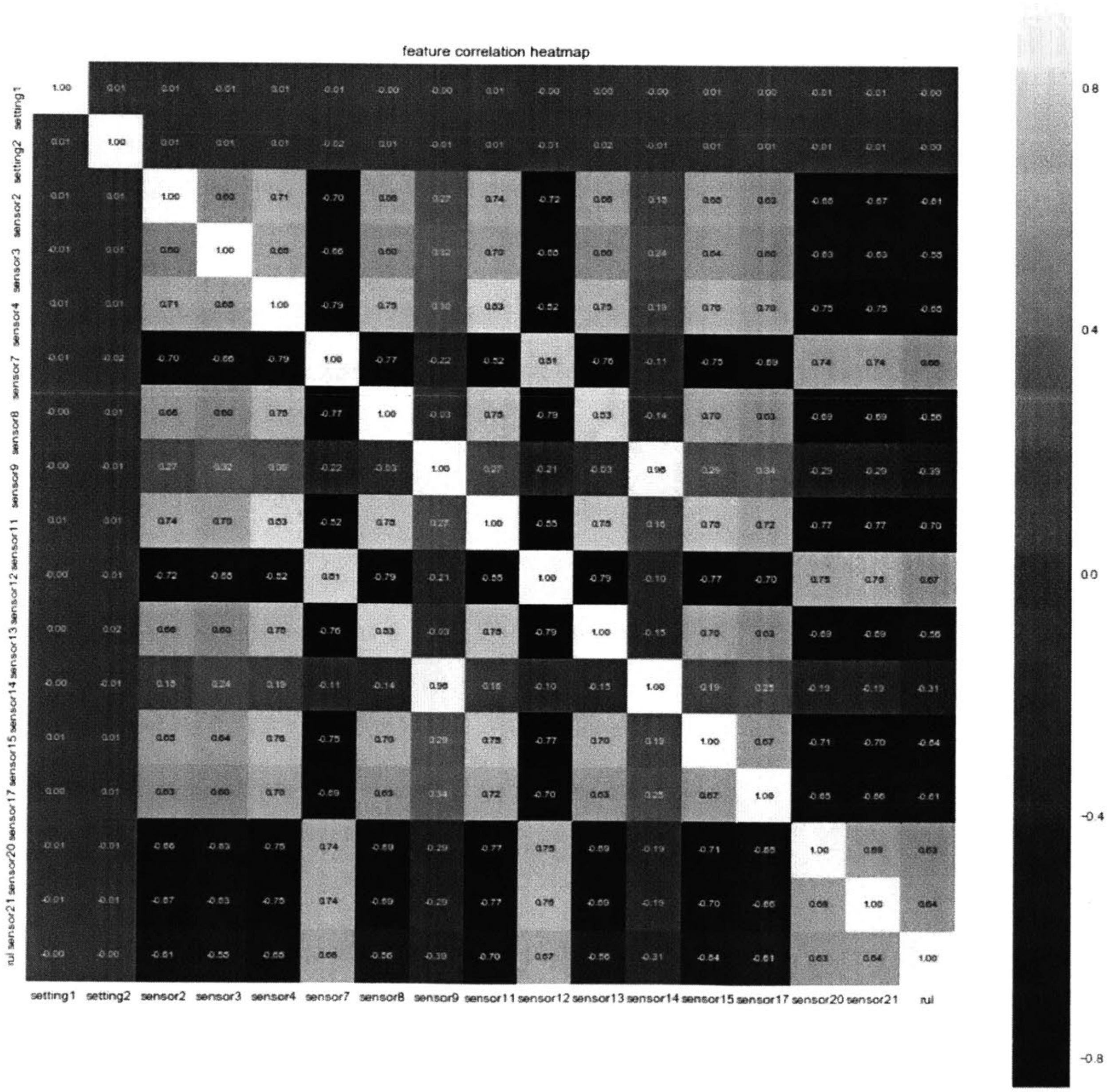

Figure 54. Feature Coefficients Weights

5.1.3 Binary Classification Model Selection and Evaluation

Whether a machine will fail within the next N cycles is predicted through classification. Binary classification models are as follows:

- Logistic regression
- Decision tree
- Random forest
- Support vector machine
- K nearest neighbor

Binary Classification performance metrics are as follow:

- Accuracy
- Precision
- Recall
- F1 score

Assuming that positive (P) means a failure, while normal (N) means no failure. The relationship is illustrated in Table 17 below.

Table 17. Relationship of Performance Metrics

		Predicted	
		P	N
Actual	P		
	N		

The equations for metrics is shown as follows [78]:

Accuracy=TP+TN/(TP+TN+FP+FN)

Recall=TP/(TP+FN)

Precision=TP/TP+FP

F1 Score=2*Precision*Recall/(Precision+Recall)

Please see the table 18 below for the comparison of Binary Classification models based on performance metrics.

Table 18. Comparison of Binary Classification Model based on Performance metrics

	LogisticRe	DecisionTr	RandomFore	SVC	KNeighbors
accuracy	0.983201	0.976405	0.986714	0.983735	0.981139
Precision	0.745614	0.532764	0.829167	0.753191	0.657993
Recall	0.512048	0.563253	0.599398	0.533133	0.533133
f1 score	0.607143	0.547584	0.695804	0.624339	0.589018

5.2 Optimal Approach

To achieve deep learning for regression, the neural network is built through Keras library. The first layer is built as a Long Short Term Memory (LSTM) layer with 100 units followed by another LSTM layer with 50 units [79] [80]. Dropout is applied after each LSTM layer to aviod overfitting. Final layer is a Dense output layer with single unit and linear activation. Please see the table 19 below for the architecture of the neural network. Similarly, this neural network architecture works for the binary classification as well. The only difference is that the final layer is a Dense output layer with single unit and sigmoid activation.

Table 19. Architecture of the Neural Network for Regression

Layer (type)	Output Shape	Param #
lstm_39 (LSTM)	(None, 50, 100)	47200
dropout_39 (Dropout)	(None, 50, 100)	0
lstm_40 (LSTM)	(None, 50)	30200
dropout_40 (Dropout)	(None, 50)	0
dense_20 (Dense)	(None, 1)	51
activation_20 (Activation)	(None, 1)	0

```
Total params: 77,451
Trainable params: 77,451
Non-trainable params: 0
```

After this optimal approach, the regression performance largely improved. Please see the table 20 below for the improvement.

Table 20. Comparison of Regression Performance between Baseline and Optimal Approach

Metrics	Baseline Approach (Linear)	Optimal Approach	Improvement Rate
Mean absolute error	37.163542	13.877025	63%
R2 score	0.327030	0.805450	146%

5.3 Future Research

There remains room for improvement, and future research are as follows:

- Models could be used for Multi-Class Classification to predict failure possibility in next 30 and 15 steps, and set different level of alarms
- Noise removal could be optimized through auto encoder neural network
- Additional feature engineering could be done such as moving average and standard deviation, change from initial value, etc
- Parameter optimization could be done through grid search

Chapter 6: Summary and Conclusion

6.1 Summary

Chapter 1 introduced the background of digital age and expectations for digital technologies' application in industrial sector. The main problem was stated and broke down into detailed questions. The corresponding answers are as follows:

- The most urgent industrial issue to be addressed is reduce unplanned breakdowns and cost, and raise productivity in production.
- The most valuable and feasible solution to this problem is industrial predictive maintenance.
- The most suitable technologies for this solution are in six stacks: device, connectivity, storage, platform, analytics, application; AI and machine learning is a key technology in this solution.
- To design a validation experiment, a detailed plan is required including goal, data, methods, steps, and tools.
- To implement this experiment, a baseline approach and an optimal approach are required, including data preparation, selection and evaluation of both regression and classification models with machine learning and deep learning approach.

Also, the primary and secondary research objectives were set. The research approach was introduced in Chapter 1.

Through the comprehensive analysis of the digital industrial ecosystem in Chapter 2, most dimensions of the digital industry were elaborated. Digital industry initiatives were explained including Industry 4.0, Industrial Internet, Made in China 2025, The Industrial Value Chain Initiative (IVI), and alternative initiatives. Through technology analysis, many significant technologies were illustrated, such as IT/OT convergence, M2M, IoT, TSN, OPC UA, Cloud Computing, Edge Computing, Big Data and Analytics, AI and Machine Learning, etc.

Then, stakeholders in three tiers (edge tier, cloud platform tier, and enterprise tier), in five domains (control domain, information domain, operation domain, application domain, and business domain), and in many technology areas were discussed.

Chapter 2 also analyzed the digital industrial market status and trend in three dimensions (vertical, product/service, and geography). Then strategies were discussed including product vs service and platform vs application (horizontal vs vertical). Various platforms provided by major companies and startups in specific verticals were analyzed and compared.

Chapter 3 illustrated the valuable and feasible solution - predictive maintenance. The work first described background of industrial maintenance, then analyzed importance, suitable scenarios, market trend, business model, and related technologies of predictive maintenance.

To implement the predictive maintenance solution, an experiment plan was designed in Chapter 4. The goal was set; the data was accessed; detailed methods and steps (baseline approach and optimal approach) were planned; and tools were determined.

Chapter 5 elaborated the implementation process of the experiment with both baseline approach and optimal approach. The work described the process of data preparation, selection and evaluation of both regression and classification models, as well as deep learning procedure through neural network building and optimization. The research turned out to be a good success. In addition, areas for improvements were discussed.

6.2 Contributions and Expectation

Through the comprehensive analysis of the digital industrial ecosystem and the valuable and feasible solution as well as design and implementation of the experiment, it is not difficult to conclude that predictive maintenance is more likely to be the breakthrough point into digital industry. With feasible solution and technology, specific domain knowledge, and accumulative experience, predictive maintenance is able to create sufficient valuc in a varicty of verticals, and benefit the virtuous circle of the digital industry.

I hope this thesis will help people to better understand what the digital industry ecosystem is, why predictive maintenance is a key solution, and how to implement predictive maintenance themselves with a step by step machine learning approach.

Also, I hope this research can help readers to clarify their thinking of digital industry from the system perspectives. These analysis and illustrations are beneficial for people to find out root-causes of problems and dilemmas in the digital industry. In this way, people are likely to have confidence in the development, adoption, and advancement of AI in the digital industry. Despite challenges, industrial AI in the digital age is growing faster and faster in the not-too-distant future.

6.3 Limitations and Future Research

Due to the limited time, there are some places that would be worth a deep dive. Future research is recommended as follows:

- The ecosystem could be decomposed into more levels to clarify the correlation among different elements.
- The V model could be used to better illustrate the entire process through system architecting, system engineering and project management.
- Extra work is needed to figure out additional reasons for selecting predictive maintenance as the critical solution.
- Additional work on the iteration in the optimal approach is expected in order to find out a more optimized algorithm that benefits predictive maintenance solution.
- Additional research on the data's physical significance could be done because this would help us better understand the data and make more reasonable actions such as feature engineering.

Preventative maintenance is used primarily to valuable and expensive industrial assets. Actually, preventative maintenance is not always the most cost-effectively solution for failure prediction of all assets. Cost tradeoffs should be taken into consideration, such as Accuracy-Gain tradeoff in failure prediction and Accuracy-Latency tradeoff in performance degradation detection, as is shown in the figure 55 below. When determining if predictive maintenance is an optimal solution for the particular asset of interest, judgment is supposed to be exercised. There are some systematic methods based on reliability-centered maintenance techniques might work for deciding whether predictive maintenance is a cost-effectively option for a specific industrial asset. Therefore, how to determine cost-effective zone of predictive maintenance and how to reduce the cost of predictive maintenance need additional research as well [73].

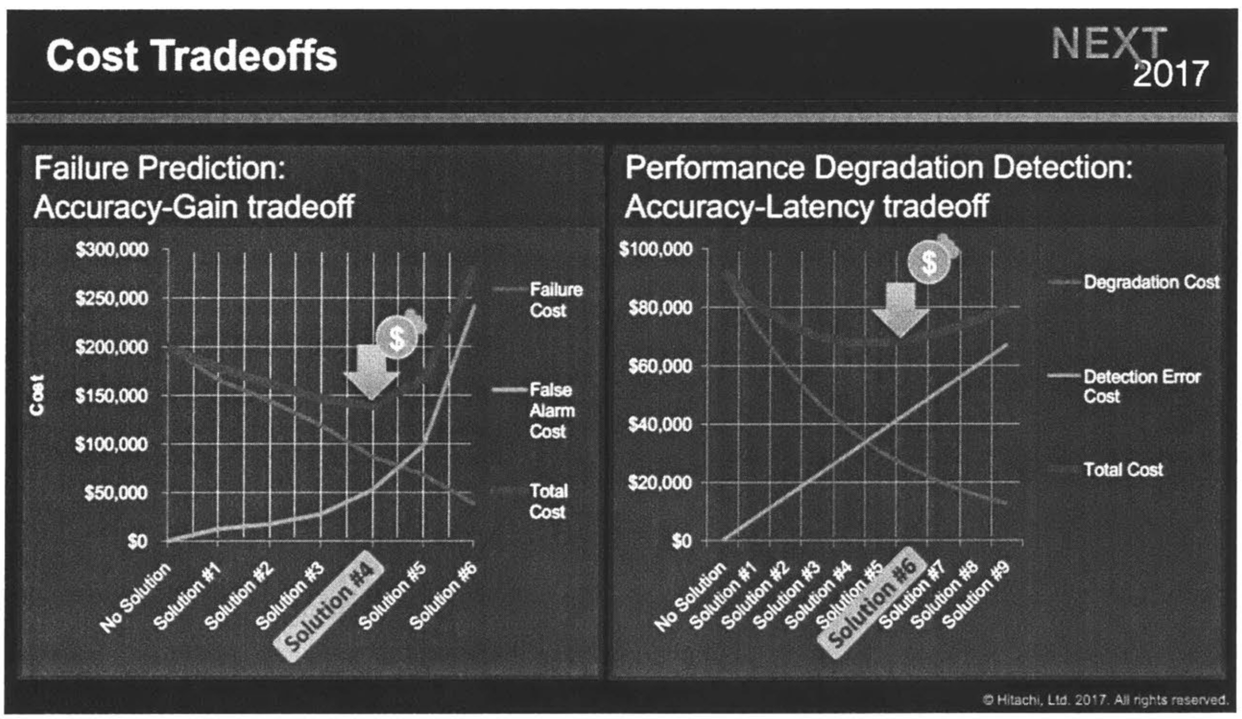

Figure 55. Cost Tradeoffs in Prediction and Detection (Source: [1])

References

[1] Chetan Gupta, Hitachi Next 2017, Lead what's next, Leveraging AI for Industrial IoT

[2] http://www.sohu.com/a/152805890_658106

[3] https://en.wikipedia.org/wiki/Industry_4.0

[4] Industrie_4_0_Cyber-Physical_Production_Systems_for_Mass_Customizations

[5] https://www.ge.com/digital/blog/everything-you-need-know-about-industrial-internet-things

[6] https://www.predix.io/resources/tutorials/tutorial-details.html?tutorial_id=1602&tag=1631&journey=Understanding%20Predix%20(for%20developers)

[7] IIC_PUB_G1_V1.80_2017-01-31 https://www.iiconsortium.org/IIRA.htm

[8] https://baike.baidu.com/item/%E4%B8%AD%E5%9B%BD%E5%88%B6%E9%80%A02025/16432644?fr=Aladdin

[9] http://www.oneroadresearch.com/made-in-china-2025-from-bulk-to-brilliance/

[10] https://www.japanindustrynews.com/2016/04/internet-things-japan-quietly-systematically-plowing-ahead/

[11] https://www.marketsandmarkets.com/Market-Reports/industry-4-market-102536746.html

[12] https://www.bcg.com/en-us/capabilities/operations/embracing-industry-4.0-rediscovering-growth.aspx

[13] https://www.mckinsey.com/business-functions/operations/our-insights/manufacturings-next-act

[14] https://en.wikipedia.org/wiki/Operational_Technology

[15] https://www.gartner.com/smarterwithgartner/when-it-and-operational-technology-converge/

[16] https://dzone.com/refcardz/getting-started-with-industrial-internet?chapter=1

[17] https://blogs.sap.com/2015/07/29/moving-from-data-overload-to-business-insight/

[18] https://en.wikipedia.org/wiki/Machine_to_machine

[19] https://www.postscapes.com/internet-of-things-protocols/

[20] https://dzone.com/refcardz/getting-started-with-industrial-internet?chapter=5

[21] https://en.wikipedia.org/wiki/Internet_of_things

[22] https://www.sam-solutions.com/blog/what-is-the-difference-between-m2m-and-iot/

[23] https://www.grandviewrescarch.com/press-release/global-industrial-internet-of-things-iiot-market

[24] http://www.basvankaam.com/2017/08/15/all-things-connected-the-i-in-the-iot-a-closer-look-part-three/

[25] https://www.automationworld.com/article/topics/industrial-internet-things/standards-and-protocols-industrial-internet-things

[26] https://www.innovationpost.it/2018/02/23/le-reti-ethernet-industriali-crescono-piu-dei-fieldbus/

[27] https://en.wikipedia.org/wiki/Time-Sensitive_Networking

[28] https://www.automationworld.com/article/industry-type/all/time-sensitive-networking-set-transform-industrial-ethernet

[29] https://www.siemens.com/innovation/en/home/pictures-of-the-future/industry-and-automation/the-future-of-manufacturing-time-sensitive-networking.html

[30] https://opcconnect.opcfoundation.org/2017/12/opc-ua-over-tsn-a-new-frontier-in-ethernet-communications/

[31] https://en.wikipedia.org/wiki/OPC_Unified_Architecture

[32] https://www.tttech.com/news-events/tech-talk/details/the-automation-pyramid/

[33] https://www.automationworld.com/automation-networks-pyramid-pillar

[34] https://en.wikipedia.org/wiki/Cloud_computing

[35] https://aws.amazon.com/what-is-cloud-computing/

[36] https://www.predix.io/community/predix-in-action/casestudy/2540

[37] https://www.ge.com/digital/predix-platform-foundation-digital-industrial-applications

[38] https://en.wikipedia.org/wiki/Edge_computing

[39] https://www.winsystems.com/cloud-fog-and-edge-computing-whats-the-difference/

[40] https://www.ge.com/digital/blog/what-edge-computing

[41] http://www.ibmbigdatahub.com/infographic/four-vs-big-data

[42] White paper: Technology and Innovation for the Future of Production: Accelerating Value Creation

[43] https://medium.com/@BonsaiAI/what-is-industrial-ai-7d8839a20df6

[44] https://springml.com/2018/03/09/automated-quality-control-manufacturing/

[45] Webinar: Enhancing Performance of Industrial Machinery with Oracle IoT Asset Monitoring
[46] https://www.slideshare.net/abstartups/dia01-08-keynotealvaromellogartner
[47] http://www.southcarolina-magazine.com/story/88360/industrial-internet-of-things-iiot-market-to-reach-12389-billion-usd-by-2021-industryarc-research.html
[48] https://www.i-scoop.eu/news/industrial-internet-things-market-2021/
[49] https://www.marketsandmarkets.com/Market-Reports/industry-4-market-102536746.html
[50] https://www.i-scoop.eu/internet-of-things-guide/iot-spending-2020/
[51] https://www.ventureoutsource.com/contract-manufacturing/industrial-internet-of-things-iot-startups-impact-manufacturing-and-supply-chains/
[52] https://www.marketsandmarkets.com/Market-Reports/industrial-internet-of-things-market-129733727.html
[53] https://www.slideshare.net/RajeshRajVarma/whats-wrong-with-indian-it-industry/12-Product_Vs_Service_Software_Product
[54] https://www.gartner.com/doc/3874883/magic-quadrant-industrial-iot-platforms
[55] http://partners.wsj.com/emerson/unlocking-performance/how-manufacturers-can-achieve-top-quartile-performance/
[56] https://www.ptc.com/en/product-lifecycle-report/iot-slashes-downtime-with-predictive-maintenance
[57] https://www2.deloitte.com/insights/us/en/focus/industry-4-0/using-predictive-technologies-for-asset-maintenance.html#endnote-2
[58] Predictive Maintenance – Market Report 2017-2022
[59] Report: Industrial Analytics 2016/2017, December 2016
[60] https://www.fiixsoftware.com/maintenance-strategies/predictive-maintenance/
[61] https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_accordion1.html
[62] https://www.researchgate.net/figure/NASA-Technology-Readiness-Levels-TRLs_fig1_228403318
[63] https://ti.arc.nasa.gov/c/6/
[64] Emmanuel Ramasso and Abhinav Saxena, Performance Benchmarking and Analysis of Prognostic Methods for CMAPSS Datasets, International Journal of Prognostics and Health Management, ISSN2153-2648, 2014 014, access date March 5, 2018

[65] Pin Lim, Chi Keong Goh, Kay Chen Tanand Partha, Estimation of Remaining Useful Life Based on Switching Kalman Filter Neural Network Ensemble, ANNUAL CONFERENCE OF THE PROGNOSTICS AND HEALTH MANAGEMENT SOCIETY 2014

[66] Tianyi Wang, Jianbo Yu, David Siegel, and Jay Lee, A Similarity-Based Prognostics Approach forRemaining Useful Life Estimation of Engineered Systems, 2008, INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT, , access date March 10, 2018

[67] Felix O. Heimes, BAE Systems, Recurrent Neural Networks for Remaining Useful Life Estimation, 2008 INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT

[68] Abhinav Saxena, Kai Goebel, Don Simon, Neil Eklund, Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation, PHM08

[69] PHM08 Prognostics Data ChallengeDataset

[70] https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2014/phmc_14_063.pdf

[71] https://datascientistinabox.com/2015/11/30/predictive-modeling-of-turbofan-degradation-part-1/

[72] https://www.quora.com/What-is-the-meaning-of-min-max-normalization

[73] https://www.fiixsoftware.com/maintenance-strategies/predictive-maintenance/

[74] https://gallery.azure.ai/Experiment/df7c518dcba7407fb855377339d6589f

[75] https://rstudio-pubs-static.s3.amazonaws.com/115270_921cf522f45242cd8e620cb3948c4c69.html#/

[76] https://www.kaggle.com/billstuart/predictive-maintenance-ml-iiot

[77] https://github.com/Samimust/predictive-maintenance/blob/master/Model%20Selection%20-%20Regression.ipynb

[78] https://www.infoq.com/articles/machine-learning-techniques-predictive-maintenance

[79] https://github.com/Azure/lstms_for_predictive_maintenance/blob/master/Deep%20Learning%20Basics%20for%20Predictive%20Maintenance.ipynb

[80] https://colab.research.google.com/drive/1tjIOud2Cc6smmvZsbl-QDBA6TLA2iEtd#scrollTo=WRfcnQH8nqII

[81] https://github.com/Azure/MachineLearningSamples-DeepLearningforPredictiveMaintenance/blob/master/Code/3_operationalization.ipynb

[82] https://github.com/prastham/iot-predictive-maintenance/blob/master/notebooks/example.ipynb

[83] https://github.com/umbertogriffo/Predictive-Maintenance-using-LSTM

[84] http://scikit-learn.org/stable/modules/model_evaluation.html#regression-metrics