CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	A case-study in the introduction of a digital-twin in a large-scale manufacturing facility
Authors	O'Sullivan, Jamie
Publication date	2020-11-02
Original Citation	O'Sullivan, J. 2020. A case-study in the introduction of a digital-twin in a large-scale manufacturing facility. MRes Thesis, University College Cork.
Type of publication	Masters thesis (Research)
Rights	© 2020, Jamie O'Sullivan. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-07 11:31:07
Item downloaded from	https://hdl.handle.net/10468/11867



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hÉireann, Corcaigh
National University of Ireland, Cork



A case-study in the introduction of a digital-twin in a large-scale manufacturing facility

Thesis presented by
Jamie O Sullivan

For the degree of
Master of Engineering Science

University College Cork
Department of Civil and Environment Engineering
Intelligent Efficiency Research Group
Head of School of Engineering:
Dr Jorge Oliveira
Supervisor(s):
Dr Dominic O Sullivan
Dr Ken Bruton

2020

Table of Contents

TABLE OF CONTENTS 2

EXECUTIVE SUMMARY 6

TABLE OF TABLES 7

TABLE OF FIGURES 8

DEFINITION OF TERMS 9

1 INTRODUCTION 10

1.1 BACKGROUND 10
1.2 OUTLINE OF THESIS 15
1.3 RESEARCH OUTPUT & NOVEL CONTRIBUTIONS 16
1.4 RESEARCH METHODOLOGY 17

2 LITERATURE REVIEW 20

2.1 CYBER PHYSICAL SYSTEMS 20
2.2 DIGITAL TWIN 23
2.3 MAINTENANCE 30
2.4 RESEARCH QUESTIONS 35

3 DIGITAL TWIN IN A LARGE SCALE MANUFACTURING FACILITY, FINDINGS & LESSONS LEARNT 39

3.1 AS-IS ASSESSMENT OF A MANUFACTURING FACILITY MATURITY FOR DIGITAL TWINS 39
3.2 REVIEW OF PREVIOUS DIGITAL TWIN ANALYSIS IN THE FACILITY 44
3.3 DIGITAL TWIN FRAMEWORK 51

4 CASE-STUDY – INTRODUCTION OF A DIGITAL TWIN 61

4.1 DT FRAMEWORK USAGE 61
4.2 PREDICTIVE MAINTENANCE PROGRAM 62
4.3 DIGITAL TWIN DATA SOURCES 62
4.4 VIBRATION ANALYSIS 63
4.5 CASE-STUDY DISCOVERIES 77
4.6 CASE-STUDY NEXT STEPS 81

5 MACHINE LEARNING & DIGITAL TWIN 91

5.1 MACHINE LEARNING INTRODUCTION 91
5.2 MACHINE LEARNING & INDUSTRY 4.0 91
5.3 MACHINE LEARNING, ALGORITHMS & FUNCTIONS 93
5.4 SUPERVISED LEARNING METHOD 96
5.5 CLASSIFICATION VERSUS REGRESSION 97
5.6 CLASSIFICATION 98

6 DISCUSSION & CONCLUSIONS 104

6.1 DISCUSSION 104

6.2 CONCLUSION 106

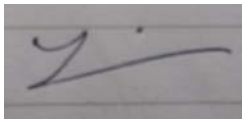
7 REFERENCES 108

APPENDIX A: JOURNAL ARTICLES REVIEWED PER SUB-HEADING 113

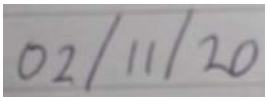
1.1 Background

Declaration of Originality

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism.



Signed: _____



Date: _____

Jamie O'Sullivan
110700855

Acknowledgments

Firstly, I would like to thank Dr Dominic O'Sullivan for his guidance and support over the last two years. I am grateful for inviting me to complete a research masters as part of the IERG group in UCC. I would also like to thank Dr Ken Bruton for his feedback and help throughout the process and for chairing the weekly feedback meetings that were so important in moving my research forward continually. Those meetings were very beneficial face to face and remotely in Year 2 due to Covid.

I would also like to thank everyone in the IERG for help during the process. The IERG was an invaluable learning environment that helped increase my learnings about my field of study and let me learn about related topics to mine that I was not directly studying. The group also improved my academic skills which needed refreshing since I joined the workforce.

Thanks also to my employers for allowing me to complete the studies and letting me use the facility as a testbed. I am grateful to everyone who gave up their time and participated in the interview process for the research. Special thank you to Eymard Gorman and Andrew Hickey who provided support as required throughout the research.

Finally, I would like to thank my family and partner for their support throughout.

Executive Summary

The exponential increase in data produced in recent times has had a profound impact in all areas of society. In the field of industrial engineering, the knowledge produced by this newly obtained data is driving business forward. Automating the process of capturing data from industrial machines, analyzing it and using the knowledge gained to make better decisions for the machines is the crux of the digital twin.

Digital twins uncover a wealth of knowledge about the physical asset they duplicate. Sensor technology, Internet of Things platforms, information and communication technology and smart analytics allow the digital twin to transform a physical asset into a connected smart item that is now part of a cyber physical system and that is far more valuable than when it existed in isolation.

The digital twin can be adopted by the maintenance engineering industry to aid in the prediction of issues before they occur thus creating value for the business. This thesis discusses the introduction of a maintenance digital twin to a large-scale manufacturing facility. Issues that hamper such work are discovered and categorized to highlight the difficulty of the practical installation of this concept. The work here highlights the difficulties when working on digital systems in manufacturing facilities and how this isn't discussed in journal articles and the disconnect between academia and industry on this topic.

To aid in the installation, a digital twin framework is created that simplifies the digital twin development process into steps that can be completed independently. Work on implementing this framework is commenced and early successes highlight the benefit of sensoring critical assets. The payback of the initial practical work is immediate, and it presents a promising outlook for the iterative development of a maintenance digital twin using the framework. The thesis' work highlights the benefit in reducing project scale and complexity and hence risk for digital systems in manufacturing facilities by following the framework developed. The later part of the thesis discusses machine learning and how this AI topic can be integrated into the digital twin to allow the digital asset to fulfill its potential.

Table of Tables

Table 1 Research methodology applied to thesis topics 18
Table 2 Asset Maintenance Strategy based on Asset Categorization 30
Table 3 Maintenance Strategy Costs [44, p. 284] 31
Table 4 Smart Manufacturing Data Project Issue Categories 40
Table 5 Success Criteria for 2016 Digital Twin Project 45
Table 6 Key maintenance issues and data types that highlight them [44, p. 8] 63
Table 7 Machinery Vibration Issues breakdown [44, p. 130] 66
Table 8 List Assets & Sensors installed in the case-study pilot phase 69
Table 9 ISO10816-1 General Machines Velocity Range Limits 74
Table 10 Centrifuge Vibration readings taken 27/02/2020 78
Table 11 Centrifuge 5 readings before and after belt replacement 79
Table 12 Option 1. Total Cost for OEM Pump Replacement 79
Table 13 Option 2. Total Cost for Local Vendor Pump Replacement 80
Table 14 Cost to Identify Issue & Complete Repairs to Centrifuges 80
Table 15 Papers reviewed 113

Table of Figures

Figure 1 Medical device industry margins (% of sales, 2006-2020) [1] 11
Figure 2 Six Pillars of Smart Manufacturing [7] 12
Figure 3 Action research methodology for phases of work 17
Figure 4 5C Architecture for CPS [25] 21
Figure 5 Three Levels of Digital Twins [37] 26
Figure 6 Progression of Maintenance Capabilities [45] 32
Figure 7 Data Input & Output Strategy from 2016 Digital Twin project 46
Figure 8 Hybrid Engineers that understand the manufacturing process and data analytics 49
Figure 9 Digital Twin framework 52
Figure 10 Extract, Transform and Load from source to storage [69] 54
Figure 11 Superposition of waves [79] 64
Figure 12 Vibration Amplitude as a function of frequency [81] 65
Figure 13 Waveform RMS Value [82] 66
Figure 14 Erbessd Phantom V Tri-Axial Vibration Sensor details 67
Figure 15 Erbessd Phantom Wireless Sensor Selection 67
Figure 16 Erbessd sensor mounted with three-axes monitoring [83] 68
Figure 17 FFT processing of a complex vibration signal [81] 71
Figure 18 FFT Waterfall Plot [84] 71
Figure 19 Shaft Imbalance [85] 72
Figure 20 Shaft Misalignment [86] 73
Figure 21 FFT spectrum for Centrifuge 5 78
Figure 22 Digital Twin Framework Version 1 83
Figure 23 Example: Sensor Selection for damaged bearing in a pump using Failure Mode Analysis Tool 85
Figure 24 Digital Twin Framework Version 2 87
Figure 25 Digital Twin Framework Version 3 89
Figure 26 Machine Learning workload is only the tip of the iceberg for data preparation work [92] 92
Figure 27 Big O Time Complexity Chart [94] 94
Figure 28 Merge Sort example [95] 95
Figure 29 Breadth First Search and Depth First Search [96] 95
Figure 30 Supervised Machine Learning workflow 96
Figure 31 Classification vs Regression 98
Figure 32 K-Nearest Neighbor Example 99
Figure 33 Diagram showing 3 Classes using K-Nearest Neighbor 99
Figure 34 Comparison of K-Nearest Neighbor Regression and Classification 100
Figure 35 Support Vector Machine Classifier [97] 100
Figure 36 SVM Hyperplane 101
Figure 37 Random Forest Classifier [98] 102
Figure 38 Graphs of ML algorithm fitting with variance and bias error [99] 102
Figure 39 Accuracy mapping due to variance and bias [99] 103

Definition of Terms

AHU	Air Handlings Units
API	Application Program Interface
BFS	Breadth first search
CBM	Condition Based Monitoring
CPS	Cyber Physical Systems
DB	Database
DFS	Depth first search
DSS	Decision Support System
DT	Digital Twin
EHS	Environmental, Health & Safety
ES	Expert System
ETL	Extraction Transformation Loading
FFT	Fast Fourier Transform
FMECA	Failure Mode Effect & Criticality Analysis
I4.0	Industry 4.0
KNN	K-Nearest Neighbor
ML	Machine Learning
MTBF	Mean Time Before Failure
MTTR	Mean Time To Repair
OEE	Overall Equipment Efficiency
OEM	Original Equipment Manufacturer
RMS	Root Mean Square
ROI	Return On Investment
SME	Subject Matter Expert
SVM	Support Vector Machine
VC	Virtual-Commissioning
WO	Work Order

1 Introduction

1.1 Background

The basis for this research stemmed from my interest in Industry 4.0 as a topic. As technology rapidly changes the world we live in, as an engineer, I want to work at the cutting edge of what is possible in manufacturing facilities. As a person I enjoy examining what is the adjacent possible in all aspects of life. This research masters allowed me to explore what is the adjacent possible in manufacturing engineering. It has also upskilled me in digital areas that are becoming more and more prevalent in an engineer's job description.

1.1.1 Disruption of global market by data

New technological innovations are continually changing the shape of the global economy. "Roughly 50% of current S&P 500 companies will be replaced over the next 10 years" [1], this rate of churn is led by new companies powered by big data analytics entering the fray. Data companies such as Facebook, Google & Netflix are disrupting their industries but also challenging the status quo in others. Historical cornerstone companies in industries are being disrupted or replaced by new start-ups that are embracing new ways to work. Innovative artificial intelligence systems are challenging traditional working methods. Data created by 2025 will be a tenfold increase from 2016 levels [2].

The manufacturing industry is no different to any other and is being affected by this shift. 33% of CEOs interviewed by KPMG in a manufacturing worldwide survey agree their organization is struggling to keep pace with technological innovation [3].

1.1.2 Manufacturing Sector

Globally manufacturing has slowed in recent years with a strong risk of a downturn [4]. The nature of modern business with its global supply chains has increased company's exposure to the impact of global catastrophic events [5]. Volatility in supply of raw materials and demand for finished goods mean an issue anywhere in the world can have a profound effect on manufacturers [5]. The changing geopolitical landscape is having a knock-on effect on the confidence of consumers. 55% of manufacturing management have stated they see a return to territorialism as a primary concern [3]. The global Covid-19 pandemic will exacerbate this issue, likely to slow the markets further. With muted growth even before Covid-19, difficulty filling critical job roles and uncertainty around tariffs and costs throughout the manufacturing value chain, the manufacturing sector is entering a period of change [4].

The sector is responding to the issues in various manners, but key tactics employed include; streamlining business to key markets and customers, creating digital agility and scalability and creating partnerships for this digital work to be completed [4]. This digital work can be new and challenging for established manufacturing companies but is necessary to help them evolve and stay resilient. Manufacturing facilities are seeing an average of 17– 20% productivity gain from embracing data led manufacturing [6].

1.1.3 Medical Devices Manufacturing

The global medical devices market is estimated to be worth $274 billion and is expected to grow on average at 4.6% annually [1]. However medical

device industry margins are on the decline and have been since 2014 [1]. This is because of product commoditization and the dynamic nature of the regulatory environment of the industry [1]. For companies to provide state-of-the-art products and prevent commoditization they must embrace data and the added benefit it can offer the customer. This data can be from a host of sources and can help the customer before, during and after surgery. To become highly data literate institutions, medical device companies must transform their relationship with data across all sectors; from customer service, to sales and manufacturing.

On the shop-floor production data must become a key ally in decision making. This process has started globally under different headings and the most recognized is Smart Manufacturing.

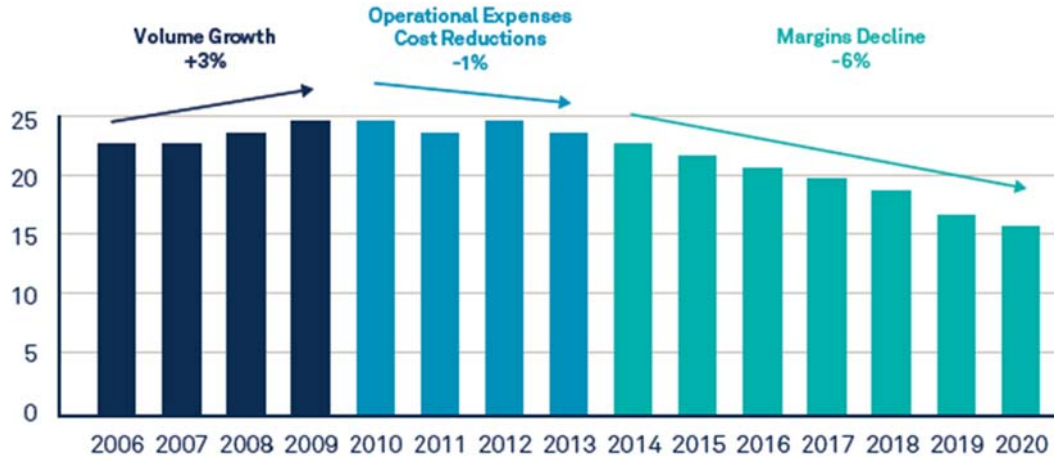


Figure 1 Medical device industry margins (% of sales, 2006-2020) [1]

1.1.4 Smart Manufacturing

Smart manufacturing is the advancements in automation, information and communication technologies in factory settings. There is not a set definition of smart manufacturing as it will mean something slightly different to various parties and will evolve over time, yesterday's new technology is todays normal equipment. However according to the National Institute of Standards and Technology smart manufacturing is fully integrated, collaborative manufacturing systems that respond in real time to meet changing demands and conditions in the factory, in the supply network and in customer needs [7].

There are several groups working in this field, the leaders being; Industry 4.0 (I4.0), The Industrial Internet Consortium, Smart Manufacturing Leadership Coalition and Technology Initiative Smart Factory [8]. The Industrial Internet Consortium goal is to accelerate the development, adoption and widespread use of interconnected machines and devices. The Smart Manufacturing Leadership Coalition is a group that leads universities, government laboratories, and research partners focused on collaborations that will transform smart manufacturing practical experiences and strategies. Technology Initiative Smart Factory group drives initiatives in smart manufacturing to increase continuous improvement, knowledge transfer and data-based decision making.

1.1.4.1 Six Pillars of Smart Manufacturing

Smart manufacturing is identified by six pillars, these are not fixed but are updated by research and development in academia and industry. The six pillars are manufacturing & technological processes, materials, data, predictive engineering, sustainability and resource sharing and networking [7].

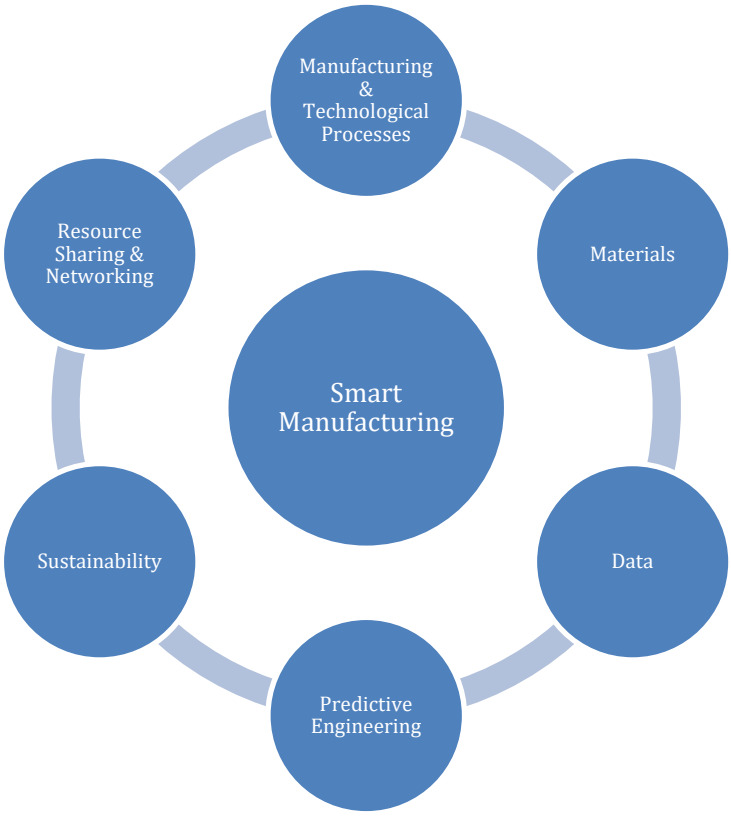


Figure 2 Six Pillars of Smart Manufacturing [7]

Manufacturing & technologies relates to the continued advancement industry has towards innovation in components and products. Additive manufacturing, adaptive manufacturing and biomanufacturing are examples of different branches of work in this pillar. The Materials pillar refers to the development of novel new materials and processes to make them, but also, linking in with sustainably to refer to the recovering of materials and products from the end of their lifecycle. This includes the possibility of landfills becoming a new source for manufacturing materials [7].

Data and predictive engineering refer to the work discussed in the remainder of this paper. Manufacturing data will be used, not only to look at the past which has been the case up to now, but to predict future cases and scenarios for manufacturing processes.

Resource sharing and networking will become an important pillar for manufacturing facilities. Previously siloed tasks and workforces will become connected digitally. For developing cyber physical systems, crucial groups involved in the work will include the operational technology (OT) team, the information technology (IT) team and the analytics team [9]. New roles will emerge as workers will need to be able to move between process engineering

and data science to be able to create and configure the correct data for manufacturing processes.

1.1.4.2 Aim of Smart Manufacturing

The aim of smart manufacturing is to improve manufacturing by turning on-site data into knowledge. Without obtaining this hidden knowledge the manufacturing industry will not develop and innovation will stagnate. This knowledge will drive better and quicker decision-making and give new insights. Smart manufacturing differs from traditional manufacturing given its focus on real-time collection and sharing of knowledge to ensure a seamless stream of operational developments. Smart manufacturing is a more capable digital application of manufacturing information where most parts of a factory are analyzed, monitored and optimized [10]. It will enable; customer centric product development, self-organization for resources and planning, self-execution for control, self-regulation for process monitoring and self-learning and adaption for quality control [11].

1.1.4.3 Smart Manufacturing Benefits

Factories adopting to smart manufacturing will see improvements in global competitiveness and reduced costs, just-in-time production capabilities and shortening production cycles, more flexibility & customization capabilities [8] [12]. They may also see a 40% reduction in cycle times, 10% improvement in time to market and 25% improvement overall operating efficiency and efficiency improvements in energy, waste and safety [13]. Companies must continue to evolve in line with these advancements in technology or risk being left behind.

1.1.4.4 Adaption to Smart Manufacturing

Three phases of adaption are required for the successful implementation of a smart manufacturing strategy [10]. Phase 1 data integration and contextualization, phase 2 simulation, modeling and analytics and phase 3 process and product innovation.

Phase 1 – data integration and contextualization. Factories must review what data is available and perform gap analysis to get a clear scope of the work required to obtain the data they need. Integrating data can be difficult, time consuming and expensive but this first step must be carried out to reap the larger rewards in later phases.

Phase 2 – simulation, modeling and analytics. Once data has been prepared it can be converted into useful information using analytics. This knowledge can improve decision-making. Data analytics far exceeds the benefits of a policy of data storage and management only.

Phase 3 – process and product innovation. Manufacturing knowledge gained from the data will aid in innovation and advancements not possible without the intelligence gained from the cyber world. These advancements will enable new business models and revenues sources [12].

1.1.5 Smart Manufacturing Concerns

1.1.5.1 Greenfield & Brownfield Sites

Concerns exist that may hamper the implementation of smart manufacturing. These challenges can depend on the facility, challenges differ

from greenfield to brownfield sites. Greenfield sites can select new technologies to install in the facility and start with a clean slate in design. Brownfield sites can have legacy issues with old systems, devices and protocols [10]. These sites may struggle with the demands of smart manufacturing and replacement of existing services may be required. Replacement of these items may not be an option for numerous reasons thus handicapping the sites smart manufacturing capabilities. Systems on brownfield sites may also be on proprietary systems or protocols. Although open standards are now available, existing facilities may be locked-in to vendor specific systems. This means only one solution may be available and hampers the facilities technological advancement. Other challenges facing facilities include upfront costs, security, workforce capabilities and leadership buy-in [10].

1.1.5.2 Cost

Cost associated with smart manufacturing works is likely to be front-loaded. Looking at the three phases of smart manufacturing implementation, phase 1 will require a significant capital investment with the payback on this not fully seen until later phases. This may be a concern to senior management as after phase 1 there will have been a large expenditure without significant returns. Companies will have to keep a view of the long-term objectives as they implement the work.

1.1.5.3 Security

Cyber-security will be vital in ensuring a smooth digitalization of industrial systems. Interested parties will be reluctant to send valuable information to the cloud if it is not secure. Also giving machines more autonomy and local control must be in the knowledge they cannot be hacked, and control changed without permission.

1.1.5.4 Workforce Skillset

Workforce knowledge and capabilities must keep pace with the advancements in smart manufacturing. Personnel will be required to understand multiple disciplines including; automation, engineering, computing, process and production. Failure to train staff and hire correctly will mean plans will not be able to be implemented due to knowledge gaps. IT, OT and analytics staff will be required to work more closely to embrace new technology and methods of smart manufacturing [9].

1.1.5.5 Business Strategy

Leadership figures at all levels of the business must buy-in to the smart manufacturing policies to ensure they are implemented properly. A clear vision and plan must be presented from senior staff to the shop floor, as works associated with smart manufacturing will disrupt normal operations. Setbacks and problems must be accepted with the objective of reaching the long-term goals.

1.1.6 Industry 4.0

The main term associated with smart manufacturing in Europe is Industry 4.0 (I4.0). I4.0 is a strategy developed by the German government to promote smart manufacturing and benefits it can add to the economy [14]. This

term is synonymous with smart manufacturing [15]. The same work globally has fallen under different local terms or initiatives such as Made in China 2025, which aims to upgrade the Chinese high-tech industry, and other initiatives in Korea and Japan [15] [16].

The term I4.0 leads on from the three previous industrial revolutions that have occurred. The first revolution (1.0) occurred in the late 1700's with the advent of power generation from steam that leads to mechanical production. The second revolution (2.0) was the industrialization of work in the early 1900's. This occurred because of the electrification of industry and the advent of mass production. The third revolution (3.0) was the automation revolution. This occurred in the 1950's & 60's & 70's due to advancements in electronics and information systems [12].

I4.0 is the fourth revolution and refers to the advancement to smart automation. Principles in I4.0 include more autonomous decision-making, the changing role of the workforce, cyber-physical systems (CPS), cyber security and advancement of IT & OT for better decision-making.

Early I4.0 goals are advancement in automation, process and production improvement and optimization. Future aims include further automation, new business models, new revenue streams through services and knowledge [12]. Arrays of new technologies are being and will be used to drive this revolution. These include advanced robotics, big data, CPS, data analytics, cyber security, artificial intelligence, machine learning, virtual reality, internet of things and additive manufacturing [7].

1.2 Outline of Thesis

This thesis investigates digital twin as a tool in a large-scale manufacturing facilities maintenance strategy. This research reviews the current literature on the topics of digital twin and maintenance in section 2 for industrial equipment. The research poses the following questions formulated from that work;

- Why are there so few case-studies of digital twins?
- What are the difficulties in constructing a maintenance digital twin?
- How would you build a maintenance digital twin?

Sections 3 provides novel insight into explaining some of the answers to the first two questions posed above. This is seen first in the results of the As-Is assessment of the facility for DTs. This is done by a structured interview process of smart manufacturing key stakeholders in the manufacturing facility. Secondly, previous DT work completed by the facility that failed is reviewed to understand the failings and prevent similar issues reoccurring. Finally, a framework is created to negate the impact of all the issues discovered. The framework aims to answer the third question posed above.

Section 4 presents the findings from the initial work in developing this framework. It highlights the early success of the work, while also highlighting challenges from this practical work in a production facility. The future state of the digital twin and its potential to leverage off machine learning technology is examined in section 5. Section 6 is the discussion and conclusions drawn from the research work.

1.3 Research output & novel contributions

The following publication was the primary dissemination of the research contained in this thesis to date:

O'Sullivan J, O'Sullivan D & Bruton K. (2021) "A case-study in the introduction of a digital twin in a large-scale smart manufacturing facility". Procedia Manufacturing Journal – Proceedings of Flexible Automation and Intelligent Manufacturing 2021. (Journal paper accepted, issuing has been delayed as the conference has been postponed until June 2021)

Below outlines the specific novel contributions of the thesis.

Section 3 Digital Twin in a large-scale manufacturing facility, findings and lessons learnt

- An As-Is assessment of the DT capabilities of a smart manufacturing facility is completed and highlights key challenges in implementing a digital solution in a production facility. These findings provide valuable insight into the gap between concept and actual installation work with DTs.
- A thorough review of previous DT project work with leading vendors is examined and lessons learnt from this work is discussed. The lessons learnt highlight the challenges even leading vendors have in this space.
- A framework for introducing maintenance DTs in a facility is drawn up. The framework provides a simplified and structured system that negates significant issues with DT installation work

Section 4 Case-study – Introduction of a Digital Twin

- The case-study work highlights the benefit of this work, while also showing the scale of the challenge of trying to understand complex data sources such as vibration data. The difficulty of combining multiple data sources is also shown by highlighting the complexity of understanding vibration data on its own.

1.4 Research Methodology

1.4.1 Active Research

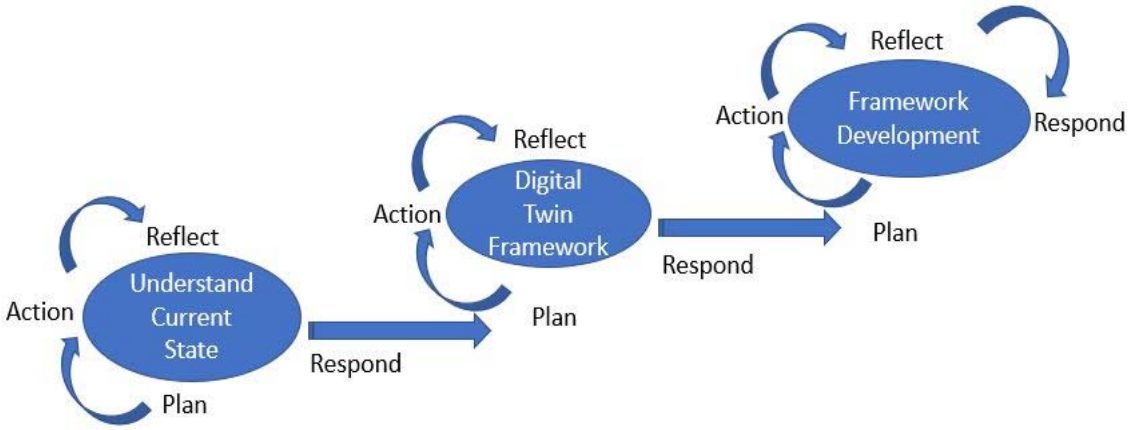


Figure 3 Action research methodology for phases of work

The research in this thesis uses an action research approach. The work involved an iterative process of; plan, action, reflect and respond [17]. This research type was chosen as it allowed a participatory nature in the ongoing workings in the manufacturing facility while developing the research topic.

This approach is appropriate "for investigating the introduction of technologies into organizations" [17]. The objective is to connect the research-based work logically with how the technology will work in practice. This approach differs to other methods such as development research or grounded theory. Active research was selected to go beyond the theoretical work of development research.

The work involves steps in each cycle; plan, action, reflect respond. It is an iterative process where once a cycle is completed the process starts again. It is participatory in nature where the researcher interacts with practitioners who wish to use the new technology and they work to take the theory and use it in practice. A plan is drawn up by the relative parties for the next phase of work, once agreed upon it is acted upon. This research style is more qualitative in nature, relying on interviews and feedback, once the process progresses and case-studies begin data can be collected and reflected upon and responded to as part of the research.

1.4.2 Research Process

The first step in this cyclical process was to interview the relevant people who were working in the field to see what work had been done around introducing a DT. This work highlighted issues around the DT introduction process. It became clear the process was a lot harder in practice than in theory. Upon reflection the response was to reduce the project scope to simplify the process.

A large amount of qualitative data collected, confirmed that a DT introduction was complex and costly. As a response a framework was developed to negate these issues. This is discussed in section 3.

The following sections go into detail on each of these cycles of action research; gathering information on the manufacturing facilities current state of

play regarding DTs, practical work around developing a DT in the facility and review of previous DT work for comparison. After that the framework produced is discussed and conclusions for future work drawn up.

The table below outlines the active research cycles for each thesis section.

Table 1 Research methodology applied to thesis topics

Plan	Action	Reflect	Respond
Understand "As-Is" conditions for installation of a maintenance digital twin	Complete structured interviews with key personnel in the manufacturing facility working in I4.0 area to gain an understanding of the "As-Is" conditions for the installation of a DT	Previous I4.0 projects had encountered a host of problems and issues that had hampered progress in the works	As part of this work it was discovered the facility had attempted to install a DT before. A full review of this project's documentation was required to gain an understanding of the necessary learnings.
Review previous digital twin work for learning	Interview the project manager for this project and review the available documentation	The scope and objectives of the project had been too ambitious and was the primary reason for the project not succeeding. The best-in-class vendors who were involved in the project could not complete the work which was a red flag	It was decided that future DTs should be built from the ground up in smaller sections that offered continuous value to the business rather than trying to construct a large single item. A framework was designed to reflect this approach
Design digital twin framework	Create a framework for the installation of a manufacturing DT for the manufacturing facility.	The framework consists of multiple parts that can be iteratively improved over an extended period as the DT develops. This works well with the slow nature of project work in regulated industries. Data collected over time becomes more valuable for machine learning	Commence practical development of the framework. Investigate physical assets in the facility that can be used for trial of adding sensors to commence DT work.
Implement DT framework in case-study	Commence practical work of	Progress in this installation work has been slow due	Continue learning and knowledge gathering around this topic by

	completing steps on the framework	regulated nature of the business. It is not a fast-moving company with regard to innovation. It is part of a large corporation and change such as advancements in I4.0 sector take time to be implemented.	investigating the future state of the maintenance DT and what it might look like. This was done by reviewing Machine Learning as a topic and seeing how it would be integrated into the DT
Understand digital twin future state with review of machine learning as a topic	Review machine learning as a topic and its relevance for use in a maintenance DT	The capabilities of machine learning can be over-stated buts its value to a DT are warranted. What is imperative to its successful use is a large amount of quality historical data	Review type and quality of data being collected by preliminary sensors.

This chapter has given an overview of the manufacturing landscape that smart manufacturing and digital twins are in. The following chapter documents the findings from the literature review on industrial digital twins and the research questions that come from this work. The proceeding chapters after that answer those questions by examining work in this space in the manufacturing facility involved in this study and by reviewing emerging capabilities in this field.

2 Literature Review

2.1 Cyber Physical Systems

A key component of smart manufacturing is CPS, these are systems to link the physical and cyber worlds to enhance manufacturing. CPS refers to a new generation of systems with integrated computational and physical capabilities that can interact with each other and with humans in real time over new modalities [18]. The functions of the CPS are the collection of real-time data from the physical world through advanced connectivity and intelligent computation, analytics and data management in the virtual world [19].

For development of smart manufacturing CPS key properties are communication, computation and control [18] [19]. The communication in the cyber space can be organized in hierarchical manner where low-level DTs feed into higher-level DTs that act as a master. When an operation occurs, the CPS can decide if the computation occurs in the physical or cloud based cyber world and how the control will be enabled. These three modes of computation and control are; physical-physical, cyber-physical & cyber-cyber. These control formats update their respective twin as the operation is being completed [19].

CPS is the integration of embedded systems with other systems using ICT. It enables the integration of horizontal and vertical enterprises in a business [20]. Vertical integration describes the connection of systems from the shop floor all the way up to business management systems. The enterprise is connected from production to resource planning. Horizontal integration is the connection along the object value chain from manufacturer to supplier to customer. There is a closed loop, which allows information to flow up and down the chain to drive product improvement.

K Alam et al propose a cloud-based CPS allowing for the system design to be inherently scalable [19]. Adoption of CPS will aid industry in production and process, engineering, material usage and waste, supply chain and lifecycle management [12]. Collection of sensory information and data from all these sources enables the discovery of insights that can drive business development.

2.1.1 Cyber Physical Production Systems

In an industrial setting, CPS in production facilities are called cyber physical production systems (CPPS), these systems could contain sensors, networks, controls, storage systems, and analytics capable of autonomously exchanging information and controlling each other and making decisions independently [12] [21] [22]. CPPS can lead to improved machine accuracy and capabilities by the integration of machine data and sensory data [23].

CPPS requires three key characteristics for the system to work; intelligence, connectedness & responsiveness [24].

- Intelligence means the system has a degree of autonomy and computation so it can take in information and make production decisions itself.
- Connectedness means the system can collaborate with other systems and all the nodes in its system. This will allow real time data exchange and decision-making.

- Responsiveness means the system has the capacity to make a production decision after there is a change in the systems environment.

Cloud computing & edge computing, sensor improvements, network capabilities and security, digital twins and other emerging technologies all make CPPS possible.

2.1.2 CPS key personnel

The increasing growth in CPS in manufacturing will require various teams onsite to collaborate so the systems can achieve their aims. Groups involved will include operational technology (OT) team, information technology (IT) team and analytics team [9]. Ensuring the engineering requirements are met are key to prevent the project running over cost, missing objectives or being cancelled [15].

OT teams ensure automation networks are running and industrial data is being obtained as required. IT teams need to develop the information systems and tools such as computing resources, software development, architecture development and data management and integration. Analytics teams must develop models and simulations that aid in decision-making. These models can be embedded in the edge at the factory floor or at a cloud level, the systems must be capable of converting data into readable information that can produce knowledge [9].

These groups working in collaboration will drive the integration of CPS. Without one team the systems put in place will break down at some level and the drive towards smart manufacturing will be more of a hindrance than an aid to the facility.

It is important that not only are these teams working towards a common goal but that they are trained up in how to create and run these new systems and informed of the benefits of implementing these new technologies. The workforce must be told why they are making changes to systems that may be satisfactorily operating currently but with the changes will exceed current performance.

2.1.3 5Cs

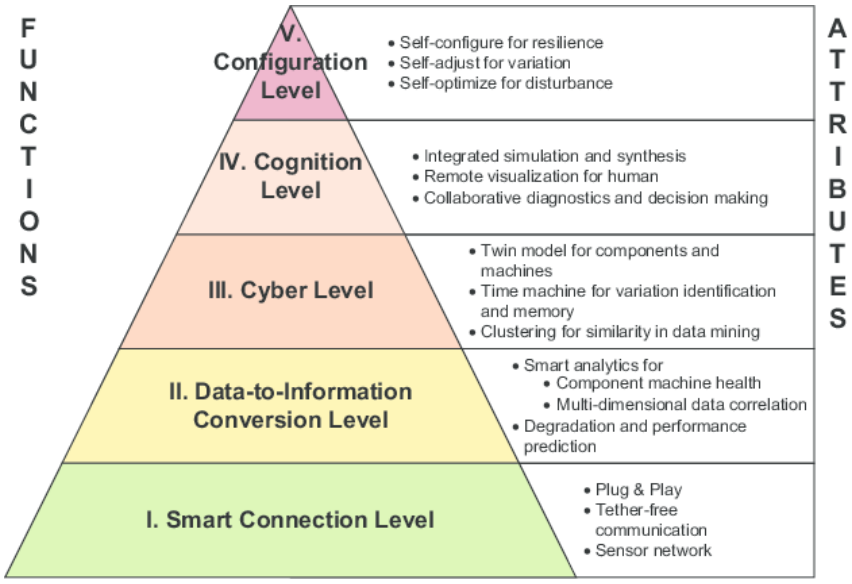


Figure 4 5C Architecture for CPS [25]

A systematic approach must be taken when developing an effective CPPS. A 5C architecture has been developed that creates a sequential workflow and illustrates how to build a CPPS. The 5C architecture for systems consists of the following [25];

- Connection – acquiring accurate and reliable data from components. Data acquisition for production.
- Conversion – extraction of information from acquired data. Inclusion of situation, status and other parameters that gave the data context.
- Cyber – carries out the changes in the cyber (digital) representation of the machine / component according to the information from the previous step and compares the new data with the previous data. Analytics of the available information.
- Cognition – generation of knowledge from obtained information. Decision making capabilities are now possible.
- Configuration – feedback from the cyber to the physical space. Application of the corrective or predictive decisions.

Stepping through parts will create a robust CPPS that increases in value to the business as each step is completed.

Sitting within the CPS as described in Figure 4 there is a virtual copy of the physical asset. This virtual twin monitors the machine components health, simulates future states and self-adjusts the machine. This virtual copy is called a digital twin and is a keystone in the smart manufacturing evolution.

2.2 Digital Twin

A digital twin (DT) is a virtual copy of a physical asset, mirroring the actions of its corresponding twin [26]. It can be used to highlight issues with the physical copy, from design through to production. It can be used to conceptualize, compare and collaborate on issues to do with the physical twin [27]. Over its lifetime it can generate a wealth of knowledge from the data it ingests and outputs.

The term Digital Twin is a relatively new one used first by the University of Michigan in a presentation to industry in 2002 [27]. The term was used in other areas, mainly the aerospace industry [14] [28] and was taken up by the I4.0 field in recent years. A recent definition of DT was a "virtual representation of a real product in the context of CPS" [29]. E.H. Glaessgen and D.S. Stargel define the DT as "an integrated multi-physics, probabilistic simulation of an as-built system, enabled by digital thread, that uses the best available models, sensor information, and input data to mirror and predict activities and performance over the life of its corresponding physical twin" [26]. The DT contains three primary components; a physical object, a virtual object and data and information connection between the physical and virtual objects [30].

The initial use of the DT was limited to helping in the analysis of monitoring the condition of the physical twin with the aim of monitoring anomalies, cracks and deformation. The aim of the DT was to change maintenance from scheduled maintenance to predicted maintenance.
The scope of the DT has widened to now include analytical optimization of the physical asset using the DT to support decision-making using models. Past and present data along with simulations predict the future, these can help interested parties make decisions about the physical or digital twin or an entire fleet of physical assets based on a digital twin(s) [31].

The DT may also provide information continuity along different phases of the lifecycle [32]. This will become key as assets are becoming better managed from design, commissioning, usage through to redesign or disposal. A DT will develop from a descriptive and diagnostic tool to being able to do predictive work. A goal will be that the DT will become prescriptive when the simulations become advanced enough.

2.2.1 Digital twin definition progression

Digital Twin still does not have a set definition. It will take time for this emerging idea to be fully formed. It must be noted though like the term Industry 4.0 it will mean something slightly different to various companies and to people in academia [32].

The DT referenced throughout this thesis refers to a DT in an industrial engineering setting. Having started in the aerospace engineering field, in an industrial setting "the DT consists of a virtual representation of a production system that is able to run on different simulation disciplines that is characterized by the synchronization between the virtual and real system, thanks to sensed data and connected smart devices, mathematical models and real time data elaboration" [32].

The above definition is satisfactory, although one could remove the word "production". This restrains DT to only production systems in a factory. The essence of smart manufacturing is the connection of all systems in the factory

setting to create a federated CPS. This would include logistics, supply chain, building services and other systems. Only when these different areas are fully integrated is the aim of smart manufacturing being fulfilled. This connected model is still some time away but with the rapid technology advancement in all areas its realization should be soon. DTs sit within this integrated CPS.

A key point in previous definitions is the connection between the physical and virtual objects. E.H. Glaessgen states that the DT is "mirroring the life of its corresponding twin" [26][ICFOIT 1], E. Negri mentions the "synchronization between the virtual and real systems" [32]. Both authors are highlighting the importance of the connection between the physical and digital states.

2.2.1.1 Digital Twin, Digital Shadow & Digital Model

W. Krizinger et al seeks further definition in relation to the connection of the twins. He defines three levels of integration; digital model, digital shadow and digital twin based on autonomy of data flow. This extra level of detail if followed will make it easier to identify the capabilities of a digital object [33]. A digital model refers to a digital object where there is manual transfer of data in both direction between the physical twin and the digital twin [33]. Digital shadows refers to a digital asset that has an automatic data flow from the physical to the digital asset and a manual data flow in the inverse direction [33]. Finally, a digital twin has an automatic data flow from the physical to the digital asset and an automatic data flow also in the inverse direction, from the digital to the physical asset [33].

Although the definition above of a digital twin calls for an automatic data flow from the DT back to the physical asset, currently for a large number of machines studied in the manufacturing facility this level of autonomy is not possible. This is also true for a lot of industrial engineering sites globally. The data still needs to flow back to a subject matter expert (SME) to make a decision for the machine. Lights out manufacturing is still a way off, and the knowledge of the SME cannot be underestimated. For this reason the term DT should allow the flow of data automatically back to the asset or back to the SME for either to make a decision.

2.2.1.2 Digital Twin in Business

A definition of DT from industry by Colin J. Parris, GE, is "a living model that drives a business outcome" [34]. This definition begins to show where the benefit is from an industry perspective compared to earlier academic definitions. Business see the DT driving business decisions and plans. The accumulation of knowledge for a business about its assets will only increase their capabilities. This definition covers the flow of data from the DT directly back to the physical twin and through a SME before they make a decision for the physical asset.

2.2.2 Digital Twin Scope

The three key components of a digital twin are the physical object, the digital object and the connection between them. This connection allows the flow of data between the two objects but also the movement of information such as actions and commands [27]. The two twins would populate a unified repository that would enable this connection [27]. Models in the DT could take the relevant

data from the repository and generate information that could feed more models or a twin.

A DT is a not a silo of all the data collected from a products design and use phases. The models in the DT are specifically designed for their intended purposes. A DT can be misunderstood as a twin suitable for all kind of tasks and this is too extensive an interpretation [35]. The DT refers to a virtual copy of a component, product, systems or process by a set of well-aligned, descriptive and executable tasks [35]. DTs are specifically designed for their intended purposes [36].

A DT in an industrial setting can vary in size depending on many factors; its objective, it's budget size or it's data pool. A DT could be a single component in a machine all the way up to a DT of a full industrial process. DTs can be split into three levels in the factory setting; unit level, system level, system of systems level [37].

2.2.2.1 Digital Twin – Unit Level

Unit level refers to a single piece of equipment. At a basic level the DT is a single digital model with connections to the physical twin. For a more complex piece of machinery the DT may be made up of various models; geometrical & spatial, behavior & functionality, usage & degradation, metrology, material properties and others [38]. These models may be the same or separate, but they all are connected back to one physical entity.

2.2.2.2 Digital Twin – System Level

The next level is the system level, which would be a cyber version of a larger process containing multiple unit level DTs. These systems are able to execute tasks for the system based on information fed back into itself from the system level DT. Rosen et al put forward a proof of concept of this system DT system [39]. Such a DT will convert sensor data to knowledge and can complete tasks autonomously.

2.2.2.3 Digital Twin – System of Systems

A DT of the "system of systems" (SoS) will give a holistic view of overall operations [37]. At this scale the DT will be of various systems on an industrial site such as production, logistics, supply chain, building services. Or the DT may be the combination of similar processes from all the global sites. This DT of a single process but in different locations around the world will allow comparison of sites and could allow autonomous feedback as required to improve sites. This knowledge collected at the macro scale can be used to run the businesses more effectively.

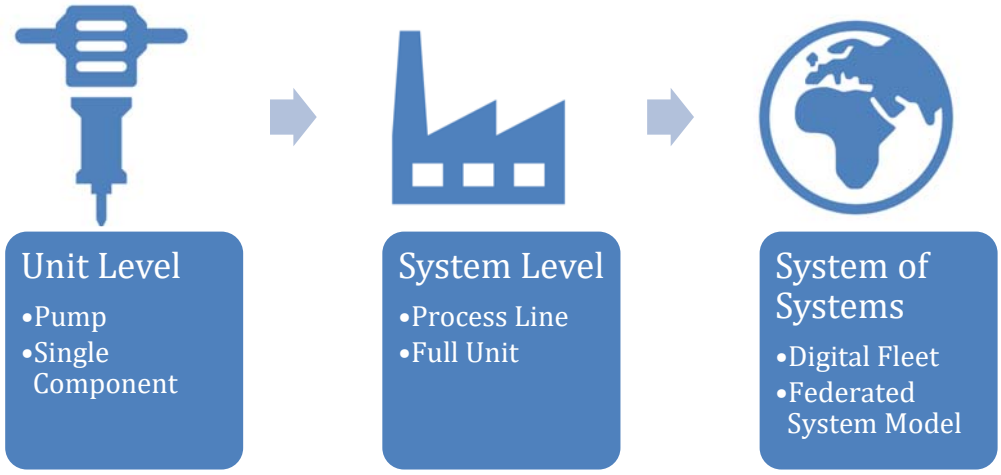


Figure 5 Three Levels of Digital Twins [37]

2.2.3 Digital Twin of a Fleet

A collection of identical DTs, at any business level; of a production part, machine or process can give invaluable insight into the physical fleet of units [35]. With the data collected from a fleet the design data can be compared with live data to look for anomalies, issues or improvements. Units in a fleet can be compared with each other from all over the world. This comparison will allow interested parties to see when maintenance is required and where units are running effectively. Knowledge from a fleet of DTs can also be fed back into the design process for the physical unit. In the current climate where global travel is being curtailed because of costs and pandemics this remote access to assets is a suitable method of work.

2.2.4 Benefits of Digital Twins

The benefit of DTs is held in the knowledge they create and how the designer uses this knowledge. DTs have the capacity to improve maintenance strategies, help with asset lifecycle management and future forecasting in various guises for the asset.

2.2.4.1 Maintenance

Maintenance of industrial equipment can represent over 30% of operating costs annually for a facility and 60-75% of a machine's lifecycle cost [35]. To be able to drive down these costs using DTs would have significant cost benefit for interested parties. DT models can predict downtime or degradation and production can be adjusted through the automation architecture [40]. The DT can help develop an anticipatory maintenance strategy increasing product quality, plant availability and reduced maintenance costs [24].

2.2.4.2 Lifecycle

DTs can help in the various lifecycle phases of an asset, in design, preproduction and production phase and decommissioning. The DT can help in

the first design iteration of an asset and it will hold even more valuable knowledge in future versions of the asset.

2.2.4.3 Prediction

DTs will have all the information of how the physical twin has reacted to an event in its environment. When this event occurs again the DT can select the best course of action for the physical twin based on how the DT reacted in its digital environment [41].

DTs will allow for the forecasting of development changes to product and processes without the need for ineffective physical mock-ups [28]. The DT can be used for the development of future systems; the DT can verify the physical objects behavior in the future systems [41]. The DT will help decrease the failure of physical twins when they are deployed, reducing expenses, time, and most importantly improve safety for users [30].

2.2.5 Digital Twin Limitations

DTs cannot solve all digital issues for a physical asset. Users must be aware of the limitations of the DT. A misconceived view is that a DT means one overall model containing all imaginable information about the physical asset. It is more likely the DT is multiple models with different purposes and information combined to give an overall picture. One DT cannot meet all purposes and tasks over a lifecycle. Interoperability, data quality and data security are some of the issues that show the limitations of the DT [15].

2.2.5.1 Interoperability

Collaboration of different companies or original equipment manufacturers (OEM) in bigger companies, must share data for the DT to be successful. These different parties will have to exchange information to make the DT work in real-time. For data analytics, the interdependent data sources must be connected in real-time to perform diagnostics that relay live information to the workers.

Interoperability Limitations

- Proprietary Systems versus open access and sharing of data
- Real-time shared access capabilities
- Real-time analytics capabilities

2.2.5.2 Data Quality

Data quality must be controlled due to the high quantity of raw data. The quality must also be guaranteed through the entire lifecycle of the DT. This is a challenging task that could limit the use of DTs. The extraction, transformation and loading (ETL) process for large sets of data will limit the DTs real-time capabilities.

2.2.5.3 Data Security

Data Security may also cap DTs because connection to the cloud for manufacturing sites can pose security concerns. DTs ability to spread information globally will be constrained by security protocols for the manufacturer. There will also be pressure on sensors to have the capability to carry out the necessary edge computation. This work may also be done on legacy systems which pose more issues as the equipment may be running on outdated controls.

2.2.6 Digital Twin Challenges

2.2.6.1 Ownership

The creation and ownership of DTs is a topic in its infancy. As manufacturers create DTs, they may be sold as part of a package with their physical twin. The manufacturer will have brought the digital object from design phase through to construction. Once the unit is bought by the user a question is then raised; who will own the data generated by the DT? Will factory owners use, supplement or replace manufacturers DTs? It is too early answer this question, but this is the type of dilemma posed by this new technology.

2.2.6.2 Privacy

There is also an issue around the data in the DT and privacy issues [15]. A manufacturer may sell a unit to different competitors. These competitors may not want their data being available to the manufacturer, as they fear competitors could get their hands on the data and get a competitive advantage. This conflict may repeat itself further down the units supply chain.

2.2.6.3 Digital Twin as a Product

Clear boundaries of what information can be used by who will have to be set out in the future. This may see a shift in manufacturers sales methods. Where previously they would have sold units as one-off sales with minimal after service. They could change their sales method to a service provider. This will allow them greater control over the DTs they produce and mean they will get feedback from the DTs for improved product development. Manufacturing with the aid of DTs may move vendors to a service-based logic, away from delivering just fixed black boxes in one-off sales [15]. The DT will be part of the CPS that will present new business models.

2.2.7 Digital Twin Lifecycle

DTs can be used through the full lifecycle of an asset, from design, production phase and decommissioning. The DT can be used to optimize quality, complete efficient and effective inspection, analytics and simulations.

B. Schleich et al, M. Macchi et al, R. Soderberg et al, M. Ayani et al and R. Rosen et al all cover DT in the lifecycle of the physical entity [28] [31] [38] [42] [39]. The concept is that the digital object in one form or another will be connected to the physical twin from design through to end of lifecycle. This closed-loop optimization allows for improvements in the physical unit's lifecycle but also for the future design of new iterations of the unit.

Terms used when writing about lifecycle vary from journal article to article, a concise version used here is; design, development, validation, maintenance and improvement [42].

2.2.7.1 Design

In the design phase stage geometric and other model types can be created of the physical twin. The digital object can be fed new data or data from similar earlier projects [38]. It is likely at this stage the digital object is only a digital model and there is no automatic transfer of data occurring between the two twins yet.

2.2.7.2 Development

In the development phase testing can be carried out on the physical or digital object. Information obtained from this testing can help to improve the final product. A detailed digital object at this stage can allow multiple testing options that might not be possible to carry out on a physical object.

2.2.7.3 Validation

During the validation phase the digital object can be used for virtual commissioning (VC). VC is the checking and verifying of systems against a virtual model [42]. The major benefits would be the reduction in costs and travel for the commissioning work. It allows some testing to be carried out prior to real commissioning on site thus helping to alleviate the pressure associated with on-site validation. The real physical controller would be tested on the DT in the VC.

Once the physical object is operational live data can transfer between the two twins. This data can be converted into valuable knowledge using data analytics and modelling.

2.2.7.4 Maintenance & Improvement

In the improvement phase the DT would be used for prescriptive analysis. Using past & present data, various simulations and models the virtual twin will autonomously make decisions that affect the physical twin. R. Rosen et al examines how a buffering, drilling and milling industrial process line can work autonomously using DTs [39]. The system DT moves the physical objects as they require different work to be carried out on them based on the data sent to the DT. Knowing how long each process takes and which will be available first the DT queues physical objects in the different stations.

Over its full lifecycle timeframe, the DT will have built up a vast silo of data. This data can used for comparison with other digital objects and for the design of future objects. It can inform future design decisions but also see how accurately design data matched actual live data. This closed loop process gives invaluable feedback which was previously not available.

2.2.8 Emulation

A key component of smart manufacturing is sustainability. When a machine reaches it end of life cycle it would previously have been disposed. An alternative is to recondition and retrofit the machine. This can be a more cost-effective solution. DTs can be used for emulation for the machine reconditioning [42]. As part of the upgrade the machine controllers must be retrofitted. The modifications, commissioning and validation to a normal machine can lead to downtime, which could mean a loss in revenue.

To mitigate against this virtual commissioning can be carried out to reduce downtime and increase final quality. As the machine systems may be old and information about them limited, they will be emulated into a DT. The difference between emulation and simulation is that emulation is connected to the real control software. The new controls can be added to this DT and virtual commissioning can be carried out. This pre on-site commissioning work can remove bugs and allow interested parties from engineers to machine associates to see how the controls are functioning. Thus, issues can be removed, and on-site commissioning time is reduced thus saving downtime and costs. When using a

DT there was a 60% drop in on-site commissioning time when reconfiguring the machine [42].

2.3 Maintenance

2.3.1 Introduction

Maintenance and operational costs contribute a large amount to the life-cycle cost of an asset and process. Exactly how much, can vary depending on multiple factors including process, maintenance strategy, product demand and asset runtime schedule, among others. These costs can be reduced by enhancing the knowledge about the assets using smart manufacturing data. Improved maintenance strategy types decrease overheads, increase machine uptime, reduce energy consumption and increase parts produced [8]. Maintenance strategies must change from reactive to proactive to realize the value offered by technology in their area.

Unforeseen downtime can be a high cost in production facilities. In one case-study a computer numerical control cutting machine (CNC) spindle failed with a cost of €25,000, failure to spot this fault caused another fault causing €250,000 worth of damage. This damage caused the production line to be shut and the downtime cost €10 million to the company [43]. A maintenance strategy implemented correctly would have raised a flag with the minor issue and repairs would have been carried out thus preventing the loss in earnings. More effective maintenance strategies leveraging smart manufacturing technologies must be used to prevent these type of breakdowns.

2.3.2 Maintenance types; run-to-failure, preventative & predictive

Depending on an asset's criticality to the manufacturing facility it can have a different maintenance strategy. General purpose assets can be run until they fail as downtime from them does not impact the facilities production. Essential assets can receive preventative maintenance, this is when the asset has scheduled interventions during the calendar year where components are replaced based on lapsed time since the last replacement. This is an improvement on run-to-failure as most components are replaced before an issue occurs. This maintenance type does not prevent breakdowns due to parts failing before the scheduled maintenance however.

Table 2 Asset Maintenance Strategy based on Asset Categorization

Maintenance Type	Asset Category
Run to failure	General Purpose
Preventative	Essential
Predictive	Critical

This issue of parts breaking before the scheduled maintenance would be detrimental to a facility if it occurred on critical assets. Hence a predictive maintenance strategy should be used on critical assets. Here the assets components are monitored and degradation to them is tracked and they are replaced based on wear-and-tear, not on pre-planned scheduled dates.

As you move from run-to-failure to preventative to predictive maintenance strategies the cost of the maintenance strategy increases as more data must be collected and tracked. Hence why not all assets have predictive capabilities. However, units breaking and failing due to less sophisticated strategies also has a cost, a cost to repair the unit and the cost of lost production time. When these factors were analyzed by V. Wowk on average the predictive maintenance strategy was the cheapest for overall cost for an asset.

Table 3 Maintenance Strategy Costs [44, p. 284]

Maintenance Type	Cost	
Run to failure	17/18	$/hp/yr
Preventative	11/13	$/hp/yr
Predictive	7/8	$/hp/yr

Table 3 above shows the cost implications of running to failure when compared with preventative and predictive. The facility in the case-study now runs some assets to failure with preventative schedules for most assets. The current maintenance workload is 60:40 split to reactive to preventative maintenance. This quantity of unplanned work is expensive from both a time and capital cost. P. O'Donovan states that in smart manufacturing "a common theme is the emphasis on transitioning operations from reactive and responsive, to predictive and preventative" [8]. This switch from reactive to proactive is what the maintenance DT will enable.

2.3.3 Maintenance Capabilities

A reactive maintenance strategy is one based on run-to-failure, a proactive strategy is where the maintenance team are fixing issues before they occur. This can be done using preventative and predictive strategies. The maintenance group will look to move from reactive work which is descriptive and diagnostic. To proactive work; which is predictive and prescriptive.

Descriptive work is explaining what happened after the event has occurred. And diagnostics is figuring out why it happened. All this work is after the event has occurred. Predictive and prescriptive capabilities allow a

maintenance team to resolve events before they occur, thus preventing failures and lost production time. Predictive work is figuring out what will happen using data and prescriptive is when the maintenance system has the capabilities to tell the workers what to do before an event happens.

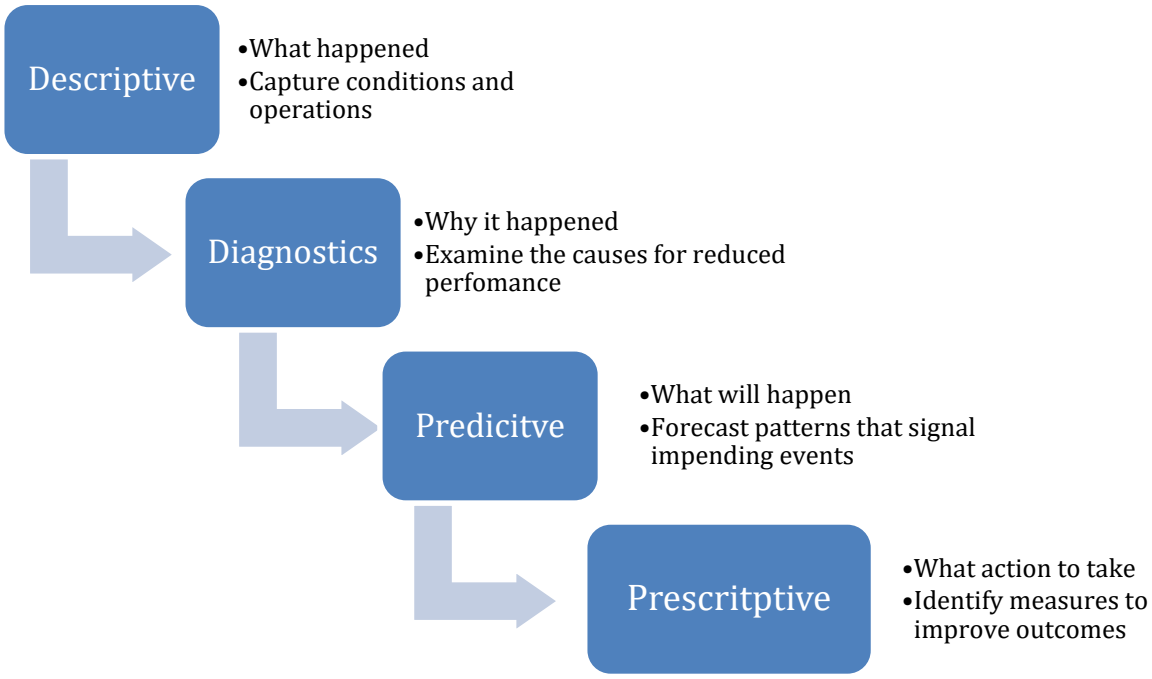


Figure 6 Progression of Maintenance Capabilities [45]

2.3.4 Smart Manufacturing and Maintenance

Maintenance costs for companies can be reduced by improved maintenance strategy capabilities. Maintenance strategies in large-scale manufacturing facilities can be made smarter using the easier access to data because of the recent advances in computing & sensor technology.

Smart maintenance strategies will be based on prognostic work, where the issue is predicted before it occurs rather than diagnostically where the work is post-event repair and analysis [46]. A.K.S. Jardine et al. proposes steps for data acquisition, data processing & maintenance decision making that fits this model [46]. Brenno Menezes et al. outlines an analytics evolution from descriptive work in the past, to a predictive framework now and cognitive analytics in the future where the analytics work is automated and adaptable to the situation for maintenance [47]. Prognostics cannot fully replace diagnostics as unplanned fault issues will always occur; diagnostics can complement a prognostics-based approach. The lessons learned from the diagnostics work can feed back into a prognostics system to improve it.

2.3.5 Advanced analytics and maintenance

Advanced analytics will be required to move a maintenance strategy from descriptive and diagnostic to predictive and prescriptive [45]. Using advanced analytics for maintenance has been discussed for many years, with the idea of using neural networks for diagnosing faults on rotating machinery discussed back in 1997 [48]. The technology and data capabilities were not in place for widespread use of advanced analytics in maintenance in the past. It is only in

recent time due to Moore's law that improved and cheaper hardware and software are available that allow teams to complete the required analytics.

The first step towards this smart maintenance is conditioned-based monitoring of assets. This allows constant collection of data from sensors for key components for critical assets.

2.3.6 Condition-based monitoring

Condition-based monitoring (CBM) of machines allows staff to proactively resolve issues before they occur and reduce unnecessary scheduled maintenance work [46]. As part of I4.0, IOT advancements and CPS development CBM is becoming more accessible for manufacturing companies. CBM can reduce maintenance costs by up to 30% and reduce machine breaks by up to 75% relative to a scheduled maintenance strategy [49]. For a manufacturing facility, maintenance programs must prevent or reduce downtime, production decreases, delays and supply chain issues by minimizing equipment failure. CBM work helps reduce these issues.

Monitoring & prognostics is an interconnected workflow with multiple linked parts. For CBM to be introduced into a manufacturing facility all the various parts would need to be in place; real-time & historical machine data, data analysis capabilities, prognostics and an overall CBM strategy. R Gao et al. and Y. Peng et al. provide overviews of this ecosystem [50] [51]. The digital twin for maintenance is one key part of this ecosystem.

2.3.7 Maintenance & Digital Twin

A DT of assets will allow maintenance groups to monitor the machinery and predict scenarios before they occur in real-time. Maintenance has been listed as the second most focused on area for DT publications behind production planning and control in one journal article [33]. In another article prognostics and health management ranked as the number one application for DTs ahead of design and production [52]. Thus, maintenance and DT are being investigated by researchers.

A DT can help reshape a factories maintenance program. The use of a DT will change the maintenance strategy from preventative to predictive. Preventative scheduled maintenance can lead to premature replacement of components and unnecessary maintenance [10]. Predictive maintenance using the knowledge from the DT optimizes replacement of components and maintenance occurs when required [10]. Prescriptive capabilities can tell workers what course of action to take based on knowledge gained through collected data, algorithms and modelling. Decision support systems and automatic failure mode effect and criticality analysis tools would be examples of prescriptive analytics.

The digital twin can offer a service to the maintenance team, it can be a new tool for decision making and the control process. Some possible solutions are listed below [53];

- Fault Diagnosis and Monitoring
- Prognostics
- Scenario Execution
- Notifications

For Fault Diagnosis and Monitoring a maintenance DT would offer real-time status for the asset's health. It would also diagnose previous issues based on historical data and knowledge added to the twin through a knowledge management system.

The Prognostics capabilities of the DT is forward looking, it is forecasting future performance of the physical twin based on the DTs outputs from sensors and simulations.

Scenario Execution is the automated changing of the physical twins' parameters based on the DTs outputs. The physical twin is making changes which impact its performance to prevent an unwanted scenario from occurring before it does.

Notifications would be a more proactive add-on to the monitoring aspect discussed above. The DT rather than making automatic changes like in the scenario execution the DT will notify SMEs or managers of the assets current status and that they must make a manual intervention for the asset.

The DT will allow the maintenance team to implement predictive and prescriptive capabilities onto key assets, thus saving money in the long run. Assets would be categorized from critical to essential to general purpose, the maintenance strategy for each then would be based on its criticality rating. Then DTs can be created for the relevant physical assets.

2.4 Research Questions

2.4.1.1 Sub-headings for articles reviewed

As part of the review process into the concept of a DT, academic literature on the topic was reviewed to get an understanding of what was being studied currently. What became apparent was that there was a lack of built industrial case studies of DTs. To investigate this finding, and with a focus towards maintenance, DT literature was reviewed thematically with relevant headings in this area. The categories were; level of integration, setting, topic, maintenance applicability and data pool.

2.4.1.2 Level of Integration

The first heading "level of integration" was selected as the term DT had different meanings to various authors. Rather than introduce a new definition, the thesis references a simple but clear definition from W. Kritzinger et al. [33]. This categorization, of papers under digital twin, digital shadow and digital model highlights where the bulk of the work into digital assets was being done. The terms were explained in Section 2.2.

2.4.1.3 Setting

The next heading for review was "setting", this referred to where the twin was being installed. Was the twin for a piece of equipment in a fully regulated industrial production floor or was the twin in an academic testbed bed facility. The setting was important as it reflects how ready the DT example is for a production floor and how applicable the knowledge from the article. This section highlighted the lack of published case-studies of installed manufacturing facility DTs.

2.4.1.4 Topic

The "topic" category showed what type of physical assets were being discussed in the papers. The topics referred to what type of equipment the physical twin was. Different departments require DTs of different assets for various reasons.

2.4.1.5 Maintenance

"Maintenance applicability" ranked the papers reviewed for their relevance to the topic of maintenance as low, medium, or high.

2.4.1.6 Data

Lastly the papers were reviewed for "data pool" to highlight where the data for the digital asset was being pulled from and how large this data set was. The "data pool" section showed that papers generally discussed data sources as a broad term that covered vast amounts of data or the data source was a small set of data points specific to the machine discussed in the paper. All the papers reviewed are tabulated under the headings in Appendix A.

2.4.2 Results

Of the papers studied only 38% referred to digital twin under W. Kritzinger et al. definition and most of these papers are discussing the theory on the topic. Some laboratory examples were built by M Vathoopan et al, S. Haag et al. and W. Yang [14] [36] [40]. These papers DTs are of simplified systems with a

low sensor count. None of the above examples are of the complexity and scale of the DTs as discussed in the concept papers by R. Gao et al., F. Tao et al. or Michael Grieves [27] [50] [54]. Here the authors are looking for the systematic linking of data sources in the cyber and physical state for assets to drive business decisions autonomously or with minimal human intervention.

2.4.3 Digital Twin still in its infancy

Digital twin as a topic is still in its early stages. No case studies of digital assets making autonomous decisions in manufacturing settings were reviewed. Using these assets to make maintenance decisions about the physical copy is not widely seen either. More digital shadows have been created than digital twins, whether the researchers realize the difference in the terms is unclear. But to a certain extent it is also irrelevant as it is far more likely digital shadows will be the norm sooner than digital twins, due to the reduced complexity and cost of digital shadows.

Half of the papers reviewed discuss digital shadows, a term coined by W. Kritzinger et al., to explain a digital asset that has an automatic data flow from the physical to the digital asset and a manual data flow in the inverse direction [33]. This is to be expected as this setup is less taxing to design and more in line with current industrial norms. Fully automated machines making changes to themselves is not standard manufacturing practice yet and this is reflected in the papers studied. Of these digital shadows a variety of machine types are discussed with CNCs and 3D printers being analyzed in a high percentage of them [23] [55] [56] [57] [58] [59] [60] [61] [62].

It is recommended that digital shadows and digital twins should be combined and discussed under the single term digital twin for simplicity, to reduce the amount of terms. Whether the data flows from the digital asset back to the physical asset through a person or autonomously what is important is that real-time and historical data is now being used to improve the asset in some way. The preliminary DTs covered in the papers will slowly be replaced by more complex systems as knowledge of the relevant fields become more well known, such as sensor technology, data analytics and maintenance.

47% of the papers are discussing the topic in theory or as a concept. The drawback of this is that these papers discuss both machinery and data as broad terms without going into the detail. Both machinery type and data are detailed topics in their own right. Only when this detail is analysed will complete DTs become more common place. It is through this detail the finer issues around the topic will be discussed. Evidence from these broad papers indicate there is great potential in developing DTs yet they fail to help researchers in explaining how to execute the work required. Papers from Y. Peng et al. and A.K.S. Jardine et al. manage to form papers that bridge this gap of a broad overview of an area while still covering detail, it is worth noting their work is in prognostics which is a more developed field [46] [51].

2.4.4 Digital Twin not on the production floor yet

A laboratory or academic setting accounted for 23% of the DTs reviewed in the papers. These have less relevance when looking to install a DT in a manufacturing facility. The DT does not need to be as thoroughly developed and tested in the laboratory compared with a regulated manufacturing facility. 30% of papers discussed digital assets in the laboratory/production space, this

referred to digital assets that were ready or could easily be deployed in a manufacturing area. Most of these papers dealt with CNCs. They did not cover maintenance as a topic and looked more at the geometry of the parts produced by the CNCs.

2.4.5 Data from a distance or up-close

The data applicability review showed that papers covered data in two manners. One group of authors skimmed over the topic and didn't go into the detail of the data source, extraction, loading and transformation processes. The second group covered the topic in far greater detail and outlined the sources and what the findings were from the data gathered. Understandably this split generally coincided with the concept papers versus the more hands-on work. The more practical studies show in their length that extraction of data for knowledge is detailed in nature. Work from K. Leahy et al. and H Han et al. cover the detail in working with datasets to ascertain maintenance knowledge [63] [64].

2.4.6 Digital Twin and Maintenance articles

The papers reviewed had varying degrees of relevance to maintenance, papers that ranked low in this heading could be focused on different areas such as lifecycle, cloud computing or geometry of parts produced by the machine. 20% of the papers reviewed ranked high, these can be split into two subsections. One set gave detail of how a digital asset for maintenance could have been created and the process steps involved such as; data acquisition, data processing and prognostics. A.K.S. Jardine et al. and Y. Peng et al. cover this in thorough detail in their papers [46] [51]. Other authors such as H. Han et al. and M. Vathoopan et al. focused in on particular equipment types and provided maintenance advice on those specific assets [40] [64]. The combination of these specific and broader maintenance papers helped to build up a framework of how any asset's digital maintenance copy could be created.

It is apparent from some papers that some of this digital asset maintenance work is being discussed under terms such as "smart maintenance" and "prognostics" [46] [50] [51] [65]. The theory behind this work is well discussed, implementation seems to be on smaller scale than the all-encompassing DTs discussed in DT theory and concept papers. This is understandable as each subsection discussed in the DT concept papers are full fields of study themselves and to go into the detail would take too much time in an academic paper. There is a lack of examples of papers showing case studies of the finer detail of larger DTs. However, depending on the digital asset, you wish to create that detail likely exists under a different field or search term such as "smart maintenance" or "digital lifecycle". So far, few authors have packaged it all together into a full DT case study paper.

2.4.7 Research Questions

This research reviews purpose is to help understand the different aspects of work completed around maintenance DTs. There is a volume of research and discussion in the area, but the majority is at a concept level, other work then is very specific in nature focusing on single components to create digital copies of these.

Concept papers are too broad in their nature to explain how to build a DT as they gloss over intricate detail. There is a gap between these two bodies of work, this is because of two reasons; one the subject is still in its infancy and secondarily the scale up of DTs to whole production systems is complex and costly. It is important to conduct more studies on combining virtual components to start to build digital assets for CPSs. A bottom-up approach of working on small DTs and combining these to achieve further insights is a better approach than a top down manner where a large data set, from combined sources is tackled in bulk.

The literature review generated the following key questions that needed answering;

- Why are there so few case-studies of digital twins?
- What are the difficulties in constructing a maintenance digital twin?
- How would you build a maintenance digital twin?

A plan was drawn up around how these could be answered with the overarching aim of seeing a DT introduced to the manufacturing facility in question. It has been previously highlighted that there are not a significant number of industrial DTs case studies in journals, but it was not explained why [33]. This thesis will explain some reasons why. The following chapter helps answer the first two research questions mentioned above, by assessing the different complexities of industrial DTs. The chapter then introduces a framework to mitigate these complexities. Chapter 4 explains the beginnings of the process of creating a DT from initiation, while chapter 5 reviews machine learning as a topic and its relevance and capabilities for use in a maintenance DT. Others before have looked at machine learning and which techniques can be introduced into a maintenance strategy before and here we review the preferred options [46, 51].

3 Digital Twin in a large scale manufacturing facility, findings & lessons learnt

3.1 As-Is Assessment of a Manufacturing Facility maturity for Digital Twins

3.1.1 Assessment of Manufacturing Facilities Digital Twin Capabilities

This section outlines the current state of play in the smart manufacturing facility in the case-study. The main assets in the manufacturing facility are Computer Numerical Controlled (CNC) machines. There are also clean-lines, packaging equipment & coordinate measurement machines also present in high numbers. There is also a full portfolio of supporting building service equipment including heating, ventilation, air conditioning, steam, compressed air, process gases, swarf systems, water and fire suppression. After reviewing the process and building services CPS' it was found there are no maintenance DTs in the manufacturing facility. No asset is feeding live data into algorithms and simulations that recommend a corrective course of action for maintenance staff based on the assets current state and projected future state.

This section examines why a maintenance DT was not in place already in the facility. The first step in this plan was to understand the current issues with live machine data projects, similar in nature to a DT. This step was taken to gain an overview of the topic in the facility. This work would prevent repetition of previous mistakes and collate knowledge for the DT framework. It is an inexpensive task that's value is clear and requires time to collect the data.

The action in this phase of work was to interview key personnel in the manufacturing facility in a structured manner. After collecting the response data and reflecting on the insights a plan could be drawn up to respond to the state-of-play of the facility with respect to DTs.

3.1.2 DT As-Is Assessment Plan

The following personnel were interviewed in a structured manner to review their involvement & knowledge into the manufacturing facilities I4.0 & DT strategy, DT capabilities, DT projects & DT challenges. The interview process involved 30 minutes to 1-hour face-to-face sessions. Each interviewee was questioned on their knowledge of the As-Is maturity of the site for DTs based on the headings just mentioned. The following people were interviewed;

Automation Manager
Automation Principal Engineer
Automation & IT Networks & Architecture Lead
Automation Projects Manager
CNC Automation Lead Engineer
Innovation Engineer
Industry 4.0 Site lead
Maintenance Lead
Machine Safety Engineer

The answers from the subject matter experts were collected and reviewed for insights. Key findings were highlighted and are presented in the following section.

3.1.3 Digital Twin As-Is assessment findings

This section outlines the findings from the interview process to assess the As-Is conditions of the facility for DTs. The issues are summarized in Table 4 under four categories; Business, Human Factor, Project and Vendor. Business refers to issues for the wider organization overseeing multiple projects at an enterprise level. Human Factor relates to issues for workers on the shop-floor. Technical issues are in the finer detail of a job due to technical complexity and new technology introduction. Vendor refers to issues the companies who have been brought into the manufacturing facility to install services have come across.

Table 4 Smart Manufacturing Data Project Issue Categories.

Issue Category	Issue
Business	1,2,3,4,
Human Factor	5,6,7,
Project	8,9,10,11
Vendor	12,13,14

Listed below are the issues in more detail;

3.1.3.1 Data project issues; Business

1. With a large-scale manufacturing facility, to connect sensors for real-time data is a capital-intensive undertaking and in a phased approach can take a long time. Also, on a data journey for a manufacturing plant, real-time machine data is not a top priority. Machine status and alarm data would be of higher importance and needs to be collected first for reliability. So, time and data criticality for the business can affect projects.
2. Business issues can influence I4.0 projects. Project cost can hamper data availability for projects. Previous CBM work was delayed due to project funding issues.
3. Previous experience in the manufacturing facility can also hamper new I4.0 projects. Preceding attempts at data connection, prediction and CBM projects that ended negatively can impact desire on the shop-floor and C-suite for new projects in this area.
4. Being in a regulated industry can also hamper work, as cyber-security in the manufacturing facility is of paramount importance it can create issues for agile data movement. This can impact the ability to send and store data in the cloud for data analytics work.

3.1.3.2 Data project issues; Human Factor

5. The human factor can also negatively impact I4.0 work. If the DT is being used for CBM reasons a certain amount of data required for maintenance is from manual entry, for example Work Order (WO) forms and WO solutions

and problem codes. Previously this data was used for hard-copy information-only and specific troubleshooting rather than digital pattern analysis. But with the emergence of data analytics and smart manufacturing, a requirement is now there for a soft copy and an effort has been made to improve the fidelity of the data.

6. Another human factor was staff turnover. This meant projects failed or faded post implementation as subject matter experts (SME) knowledge left the group when a key member left the project group or business. Although the DT is largely self-sufficient once running, as discussed previously lights-out manufacturing is not obtainable in the medium term. So, the retention of the SME knowledge is key. This should be done using expert systems or information management systems. An expert system would be a computing system which embodies organized knowledge about the specific area of human expertise [66].

7. Workers complete tasks and actions that are not fully documented, this is what is referred to below as Hidden Factory. When trying to automate these tasks such as with a DT, these issues pose challenges that are not fully understood when the project was drawn up. Hidden Factory is also present in machine data and only in-depth knowledge of a process allows one to understand the data fully. The human who is working on the shop-floor, their knowledge is critical to creating DTs that accurately reflect the physical twin.

3.1.3.3 Data project issues; Project

8. Data quality of machine data is varied. A recent artificial intelligence (AI) project undertaken suffered from data quality issues. The machine data given to the AI specialist had inconsistencies which meant predictive work was not possible until the data was cleaned.

9. Hidden factory complexity of I4.0 projects can delay projects. Most I4.0 work is installing new technology in an industrial setting and this can pose unforeseeable issues that need to be resolved as the projects progresses.

10. Complexity of processes can also make installation of DTs difficult. Manufacturers prefer technology to be proven before introducing it into the manufacturing facility. The number of industrial case-studies of DTs is still low which indicates there are still roadblocks to its implementation [33].

11. Trying to create DTs of assets not manufactured by the company can present a whole host of problems. The time and cost required to create the DT can increase as the team seeks to understand the finer detail of the asset that they have not manufactured but do own. In previous years understanding of how to keep the asset running was enough. But to create a virtual copy of an asset across multiple model types is a significant undertaking and the work must be weighed up against the benefits.

3.1.3.4 Data project issues; Vendor

12. Buying multiple machines from different vendors to be installed in a single process line means the data quality for the system level DT should get consistent quality data across all equipment items. Again, on the shop-floor this has not been the case and has even been hard to implement on new projects. To standardize the data extraction for various assets increases the project cost by an amount and has not been accepted so far in certain projects.

13. Process and data knowledge specific to the business is a key reason why vendors have struggled to succeed in data-based projects. The

learning curve to understand the process and meaning of all the data is steep, time consuming and costly. This can hamper vendors as they try to fully understand the process workings while working with the data.

14. Some vendors manufacturing bespoke equipment are using outdated automation equipment for data extraction. This makes the ETL process, key for DT work, more difficult and time consuming. ETL is a key component of a DT and this shows the difference in difficulty between theoretical and practical implementation of a DT. Although the vendor can be requested to update their controls to allow better integration to the automation network this too can create issues. The vendor is now using new equipment which adds unwanted risk to a project. The vendor also increases the project price which means the new controls can be excluded from the project on financial grounds. Some vendors with a large customer base that are not interested in advanced automation can refuse to agree to higher automation specifications. The vendors have a dominant position in the market and do not see the need to update.

3.1.4 Manufacturing facilities maturity

The above section shows there are a host of issues when trying to complete data connectivity work in a manufacturing setting. Table 4 shows the variety of issues one must contend with for a DT to be successfully introduced into a manufacturing facility.

The manufacturing facility in the case-study is still in its early stages of its smart manufacturing journey, it is moving from level 1 to level 2 in the CPS 5C Architecture [25]. The manufacturing facility has a program of projects for real time data connection. The program work is split into three sections, monitor, predict & self-learn. The facility is still moving from step one, monitor, to step two predict. This means they are still in the process of creating robust real-time machine data connections for a host of sensors that would create enough information to allow engineers to understand holistically the health of an asset through sensory data alone.

The manufacturing facility also completed a digital maturity assessment with a leading consultancy firm and the results for connected assets section stated the facility had medium to good capabilities. The company was moving in the right direction but still had a body of work to do before fully connected assets would be achieved. This again is in-line with the findings above.

The prediction work they are commencing is directly in line with the development of DTs. It is important that the facility has enough quality data to complete this prediction work.

3.1.5 Next Steps

After reviewing and analyzing these issues, it was important that the DT development work continued without repeating errors from such previous work. The first step was to initiate further study into previous DT work that the facility carried out, step two was to develop a framework that would reduce the risk of the issues just discussed occurring again and thereby jeopardizing project success. The final step was to start implementing some tasks in that DT framework.

The following section investigates in more detail previous DT work the facility carried out. It came to light in the assessment process that a body of work had been carried out in this space before, it had not had a positive outcome as a

project and hence presented valuable lessons learned that needed to be uncovered.

With the issues raised above across a range of projects and the learnings directly from the previous DT work added to them, it was imperative that a structure be put in place moving forward. That plan is the DT framework that is discussed later in this chapter. This framework is very important in that it allows one to slot smaller tasks and projects into an overall plan to move towards DTs in the facility. The following section discusses in detail the previous attempted installation of a DT on-site, and the lessons learned from the process.

3.2 Review of previous Digital Twin analysis in the facility

Before progressing onto the framework developed for DTs in the facility. It is important to examine the previous DT work completed by the company in 2016. All documentation from the previous project work was reviewed so as not to repeat the same mistakes. The project manager was also interviewed to understand firsthand how the project progressed.

3.2.1 Overview of previous Digital Twin project

The 2016 project team consisted of a cross-department team that included shop-floor and enterprise support. The team included automation, data analytics, maintenance, machining, and digital systems engineers. The team engaged with three vendors to generate the DT.

The project was researched, scoped and had project governance like any standard commercial project. The scale of what it was trying to achieve was very significant for the business globally and if it had succeeded, it would have led to changes in a whole host of business areas. The project would have changed the make-up of the maintenance, reliability, process and operations teams if it had been a success. The information generated from the DT would have meant integrated visibility of the assets information for all sectors.

The project ultimately was not a success, this in part was because it was too ambitious with what it tried to achieve. These aims, combined with the issues like what were discussed in the previous section ultimately led to the project being shelved.

This does not mean the work offered no value, it simply highlighted the complexity, planning required, cost and need to contain expectations in relation to DTs.

3.2.2 Goals & Objectives of the Project

The goal of the project was to develop a DT for CNCs. This DT would extract and present data from the CNCs, it would measure performance, predict efficiency and potentially self-learn and correct itself. The benefit of this to the business would be significant, as to achieve those objectives would have represented a phase shift in the company's business model.
The list of benefits from a successful implementation of the DT included;

Through the DT operation, improvements to;

- Machine Output
- Downtime
- OEE
- Manual Output per person

Through DT simulations;

- Reduced Deployment Time
- Reduced Validation
- Reduced Set Up Times
- Reduced Machine Crashes
- Design Anywhere Build Anywhere capabilities
- Remote Programing of Machines

3.2 Review of pr

Success criteria for the project are listed below in Table 5.

Table 5 Success Criteria for 2016 Digital Twin Project

	Success Criteria	Measurement	Current	Target
1	As measured by standard time (labour hrs/good parts), currently ranging from .68 to .77 hours or ~15% variation	Statement	No Visibility	Standard time measure
2	An improvement of 10% is worth $1.5M/year across 2 sites	Statement	No Visibility	Standard Cost measure
3	Create Digital Twin that blends machine data with other operational sources	Implement an operational tool with ability to blend all of identified data sources. Tool operational by x date. Tool should be accessible via web link.		
4	Provide visualizations to compare across 2 sites HaaS CNC machines on:			
	a) OEE	Produce and validate OEE figures for all assets in the Pilot.		
	I. Cycle time	Produce a report that provides the best/worst cycle times per product in	No Visibility	Get Accurate Visibility
	II. Scrap	Validate accurate figures vs MES. Ability Select Family level and compare scrap figures at product code level/Machine/Site	No Visibility	0.02%
	III. Uptime	Validate accurate figures. Compare sites, machines and products being run.	No Visibility	Visible and clear comparison available
	b) SPC	Identify control limits around a machine input to control a process output.	No Visibility	Provides insights in process performance and variability
	c) Output	Identify parts produced from MES for the output reports when batches are booked off.	No Visibility	Visible Data
5	Identify and compare best/worst performers	Produce a report that identifies best/worst performers	Limited Availability from Tracksys - site level	Ability to reference and trend data across both sites in the one tool.
6	Analysis for root causes of differences between machines/sites	Analysis for root causes of differences between machines/sites add comment	No analysis currently	Automated response to queries with tangible root causes
7	Capture actionable results for the Value Stream that will support reducing differences in performance and impacting Standard Time	Identify top 5 potential opportunities to impact standard time delta supported by the data	No Visibility	5 Opportunities ID
8	Investigate relationship between OEE and Standard Time	Identify the impact a 10% increase in OEE has on standard time	No Known	Relationship known
9	Capture notable findings for additional use cases, priorities, and considerations for scaling out across business	Produce a report by x date that provides tangible data results	No visibility	Value Analysis ID for GO roll Out
10	Perform a Value Assessment on Digital Twin based analytics	Produce a report providing the investment vs return ratio identified	None Present	Report Available containing investment versus return

The scope of the project was not just CNC machine data, it included other data sources to create an overall visibility for the equipment as written about in concept papers. Data sources included;

- Site Historical data
- Workflow data
- Operations Management Systems data
- Equipment Maintenance data
- Building Management Systems data
- Tooling data

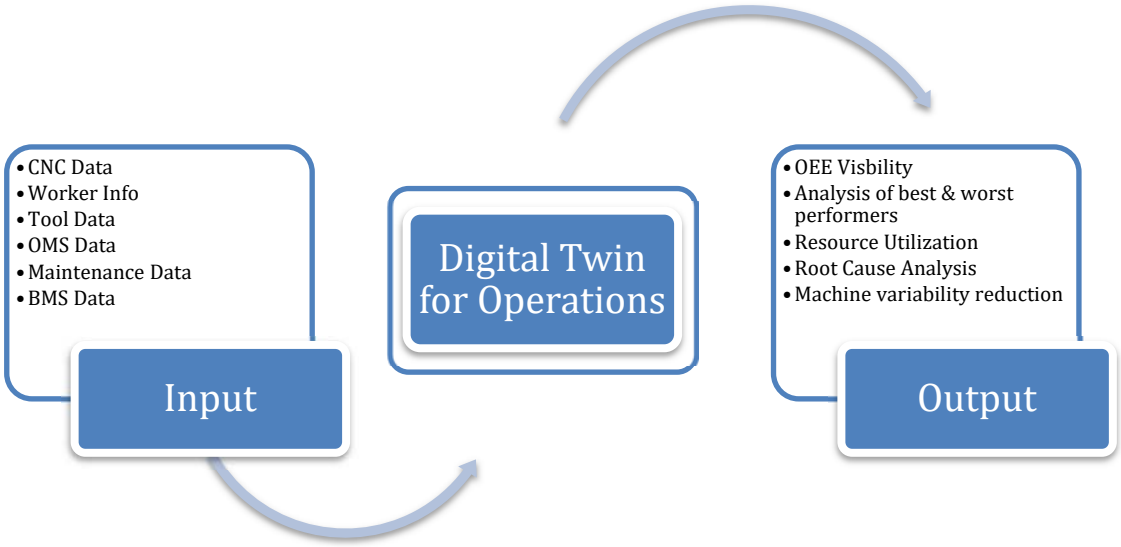


Figure 7 Data Input & Output Strategy from 2016 Digital Twin project

The goals were ambitious and grand, to step back and assess if the project was successful it would have changed, if not replaced whole departments by digitizing multiple processes currently carried out by people. When this project was carried out multiple departments such as maintenance, reliability, process and operations teams would have examined their relevant data in excels or were slowly migrating to automatic dashboards. The DT proposed to combine all these data sources into a single integrated location which would have been the DT. Steps are still being made today to complete this work such is the benefit, yet the scale is very large as data for all departments grows continually. The goal of attempting to integrate all this data in one project while introducing analytics and machine learning is highly complex and a huge undertaking. The scale of this work was likely underestimated because of the marketing from industry and academia concept papers saying this was the new norm that was about to arrive.

The project scope was scaled down when it began as the list of data sources above was too large. Once work began on the scaled down project however they still encountered difficulties.

3.2.3 Issues that arose in project-work

Once the project commenced the project team drew up a list of preferred vendors. The list was reviewed, and three vendors were contacted. All vendors stated they had expertise in the field of data and digital twins. The project team included each of the three vendors as all three had a unique skillset they could bring to the project. One vendor was a leading Silicon Valley data & digital twin specialist for all fields, one was a global leader in manufacturing & automation and the final vendor was a global technology company with experience manufacturing digital twins. Each vendor was presented with the Scope of Works documents based on the goals and objectives mentioned above.

3.2.3.1 Vendor Problems

All vendors met with the project team and after a review were given a test and learn phase of the project to allow them to implement their solution. Once all three test and learns were completed the project team was to select a single preferred vendor to roll out the solution to all facilities.

As all vendors came on-site and attempted to implement their DTs, they encountered all the issues mentioned in the previous section. The project manager stated the vendors struggled with domain knowledge, accessing the data & creating knowledge from the data. It also worth noting all issues were not on the vendors side, the full portfolio of data the vendors requested was not available to them due to data quality and gap issues.

3.2.3.2 Data Quality Issues

As the project was ongoing it became clear to the project team that clean data of sufficient quality was not present, and the vendors could not create the overarching DT desired even with the data available. The true cost of what the team looked to achieve was larger than first anticipated. For this capital cost reason, time restraints and lack of progress from the vendors the project was stopped.

3.2.4 Learnings

This project provides an invaluable insight for the company, manufacturing industry and other industries of the gap between envisioned and actual workload and the complexity of actual data projects versus their conceptual ideas. There are key learnings that can be taken from it;

- Data projects are more complex than the conceptual ideas
- Data projects are expensive if the scope is not kept in check
- Data projects require domain knowledge

3.2.4.1 Data Projects Complexity

From the literature review and what can be seen from this project is that parties can underestimate the difficulty of the ETL process for manufacturing data. Concept papers are making statements about tasks for creating DTs for manufacturing processes that require larger than expected time & financial inputs when actually carried out.

The vendors, in the case of this project, who were all industry leaders, underestimated how difficult it would be to create the DT. Reflecting now and

having discussed with the project manager this error seems to stem from the lack of awareness of the poor standard of manufacturing data compared with other industries. Each vendor struggled to extract the data, they lacked the domain knowledge about the equipment used in the manufacturing facilities industry. Even when they did manage to extract it they grappled with understanding the meaning from all the data sources. There are multiple departments in regulated manufacturing industries that understand their silo of data. To collect all this data and understand it is a complex task that has multiple stakeholders. The vendors underestimated the complexity of this work.

3.2.4.2 Data Projects can be expensive

Once the vendors began to dig into this DT project, they realized the complexity of creating the DT was bigger than they first thought. This caused the project cost to increase considerably. The value of data projects in general are only seen later in the projects cycle and expensive preliminary work needs to be put in place to ensure a strong foundation of constant data quality.

For this project, upon review the cost began to dwarf the benefits and the project manager stated these costs were likely to grow again as they moved deeper into the work. The important learning is to be aware of the cost of DT work and does the benefit outweigh it. This point can be missed in the drive to introduce the newest technology to facilities.

3.2.4.3 Data projects require domain knowledge

Lack of expertise is restricting the installation of I4.0 technology projects in manufacturing facilities [67]. While the technology transformation of smart manufacturing is much discussed the transformation and training of workers is also vital to project success [8]. Hybrid workers must have an understanding of analytics, engineering, computing, design, planning and automation [8].

For this project when the vendors began to work on the shop-floor to extract data from the physical asset to the DT it became clear to the project team that the vendors did not have the domain knowledge to do this. All manufacturing facilities are unique in their technology make up, partly due to their expansion and adding technology to legacy equipment. This creates a unique challenge for vendors when coming on-site. To bridge this domain knowledge gap for the vendor it is proposed that newly trained hybrid workers are introduced to the workforce.

3.2.4.4 Hybrid workforce

Currently there is a gap between the department engineers and the data scientists. The department engineer understands the process and collects the process data. However, they are not proficient in ETL and coding languages. They cannot manipulate data for analytics. The data scientist does not understand what the data means but they can manipulate it for dashboard, analytics and machine learning.

For the manufacturing facility to thrive in a digital age, to be able to build DTs and for vendors to come up to speed quickly when they come on-site hybrid engineers are required that can speak to both the department engineer and the data scientist. These workers need to be aware of the manufacturing processes and able to understand the language of the data scientists. Without this contact

in the business, vendors coming onsite can sink most of their time into trying to gain domain knowledge about the facilities make-up rather than upgrading it and completing their work.

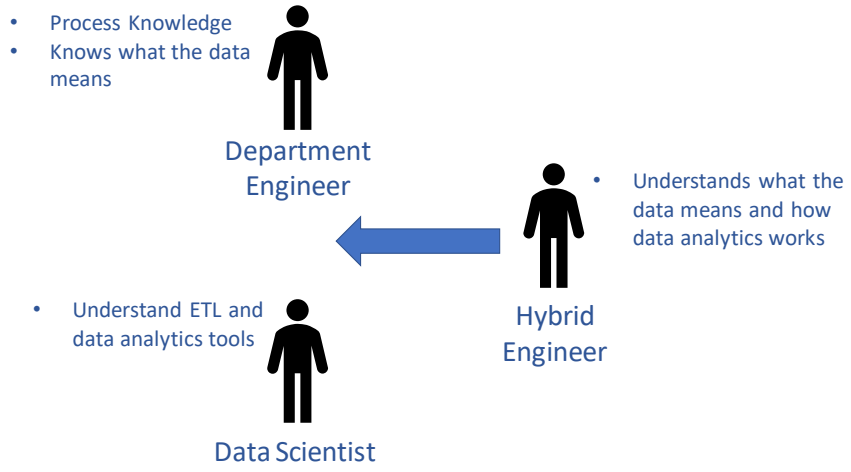


Figure 8 Hybrid Engineers that understand the manufacturing process and data analytics

3.2.5 Summary of findings reflecting on 2016 Digital Twin project

From studying this project, it is clear the aim of it was ambitious, admirable and reflective of what the manufacturing facility had been hearing from industry vendors. The project attempted to complete a game-changing digitization of processes by creating DTs of CNCs on-site that would complete many of that tasks people had suggested in academia and industry DTs could complete.

3.2.5.1 Scale, complexity and cost

The project failed because of its scale, complexity and costs trying to meet its goals and objectives. It highlighted that only workers in the facility can know their process in-depth. To try to digitize these processes by creating DTs, the facility cannot rely solely on vendors, the facility must heavily invest in data quality and systems to allow domain knowledge sharing with outsiders. It also important to upskill workers so that they understand the process and data analytics.

Based on this a framework was developed that shows how a DT for the manufacturing facility would be created. A scaled down project scope was recommended when completing this work. This reduced project size is to reduce project cost and complexity. These smaller projects sit on different steps on the DT framework and represent less risk to the business, yet each offers value by itself. Once added together over time the smaller projects combined begin to represent a DT.

3.2.6 Importance of findings

The findings from the As-IS Assessment and review of the previous Digital Twin project work highlight the value of knowledge and experience from the shopfloor. Some other authors have highlighted potential issues with DTs in

manufacturing facilities such as lag in data collection technology capabilities, unresolved cloud based analytical issues, such as network unavailability, overfull bandwidth, and latency issues [11]. As part of the literature review this commentary highlighting issues has been limited in the discussion on the DT topic. This is why the insights from SMEs highlighted from the work above is so valuable. The practical implementation of DTs can become complicated and costly in regulated manufacturing facilities and efforts must be made to prevent this from happening. Therefore, the framework is introduced in the following chapter to aid this.

3.3 Digital Twin Framework

3.3.1 Introduction

After review of previous digital work in the facility and the issues associated with data I4.0 projects it became apparent that a body of work was necessary to prevent the repetition of the same mistakes. Time and monetary expense had been incurred in the previous work that would keep repeating unless a digital data-based system was put in place.

The main issues with previous work were poor data quality, complexity of the work undertaken, scale of the attempted projects and knowledge management. With these issues in mind a review of the topic landscape with the intention of putting a system in place to negate these issues was undertaken.

3.3.1.1 Predictive Maintenance Strategy

A framework was drawn-up to help the facility when developing a digital twin. The framework helps address the problems discussed. The framework was developed after reviewing J. Soldatos et al. technical architecture for a predictive maintenance strategy [68]. The architecture covered the layers required from the shop-floor to the enterprise level for fully connected I4.0 maintenance process.

J. Soldatos et al. proposed a strategy that provided solutions at the component, machine and system level. The system targeted maintenance resolutions for prediction, diagnostics, prevention, estimation, management, remediation, synchronization and safety for assets [68]. The framework discussed in this thesis could slot into the wider maintenance strategy solution covered by J. Soldatos et al. Their work is a cognitive CPPS service development solution.

3.3.1.2 Framework make-up

The framework is made-up of multiple sections that add value to the business as stand-alone items. This is key as it allows the sections to be developed individually in a modular manner. This helps ease pressure on the building of the digital twin, as value can be shown to the business as each section is built, this means management do not have to wait until the whole DT is built to see some return. This can be a big problem with data projects as the investment is largest in the early phases and the value is not seen until later. By breaking the project down, the DT work shows its merit sooner and confidence and backing of the project persist as it progresses.

The framework developed in this thesis has a feedback loop that allows fine-tuning and improvement of the process. This is important in this digital age where change is continuous and releasing a product with no feedback is poor design.

The framework allows the creation of a maintenance DT that can act as a data access point for the physical asset it copies. The framework discussed in the following section describes a DT for aiding a maintenance team. Minor changes to certain sections would apply if the DT had a different use, however most of the structure is still relevant.

3.3.2 Digital Twin Framework Steps

The framework developed for the digital twin is shown below in Figure 9. It was introduced to reduce the impact of the issues previously discussed. The framework is split into several sections, and the remainder of this chapter goes into further detail for each section. All the sections offer benefits to the business, combined and executed properly they allow data flow between the physical and digital copies and create new knowledge for workers about the physical asset's maintenance health.

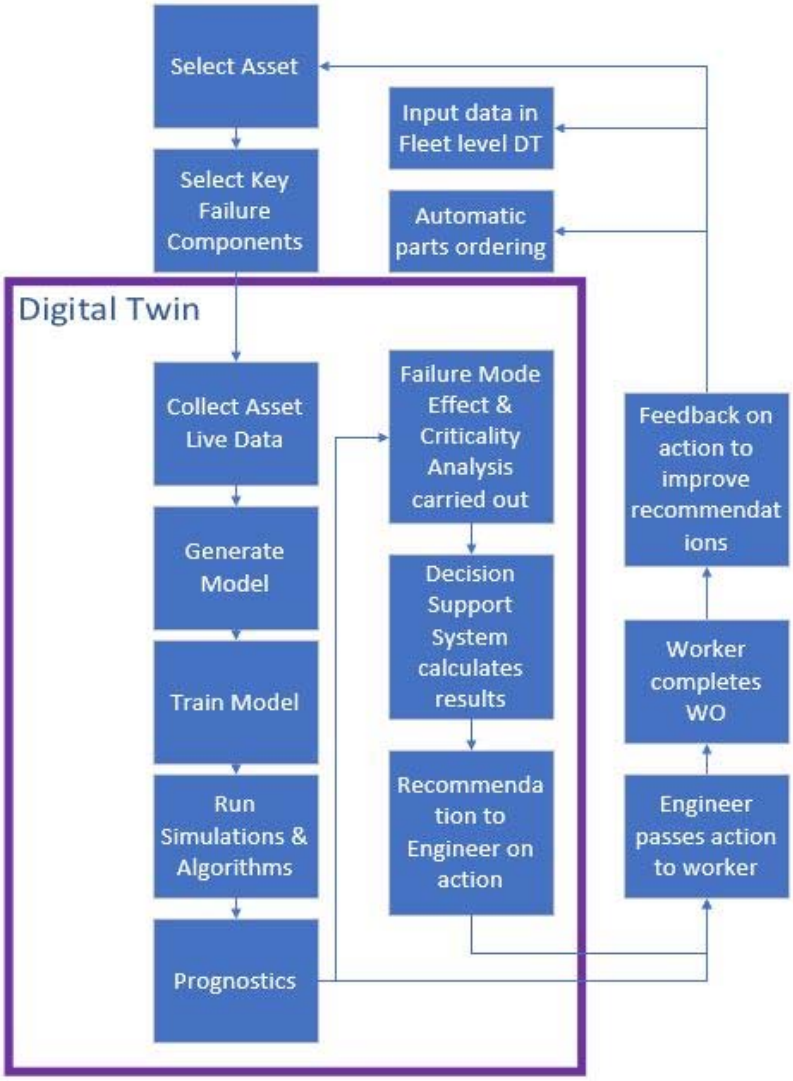


Figure 9 Digital Twin framework

3.3.2.1 Select Asset

Like any project macro planning at the start is important to project success, this step is added to the framework to ensure adequate planning takes place at the start when it is most beneficial. All large manufacturing facilities will have hundreds of assets, so selecting the correct asset to create a DT of is important. A pre-agreed selection criteria process should be followed by the team.

This should review assets by criticality, age, capital cost, consumable cost, maintenance cost & EHS issues. By creating a Pugh matrix about what criteria are most important and rating the assets the team can select the most relevant

asset for creation of a DT. This process will mean the DT will create the most value for the business. With the most pertinent asset selected the buy-in from the business into the works will be higher. This process was used by the key stakeholders in the case-study to follow.

3.3.2.2 Select Key Failure Components

Now that an asset has been selected, the team must choose the relevant components for which digital copies should be created. This step is important as it can break complex machinery down to their basic components. Here first principles of engineering can apply, so for certain components they can only fault in a set number of ways, hence reducing and highlighting the applicable sensors.

Software can be used to highlight the key failure modes for each component in the machinery. This could be cross referenced against maintenance records to show repeat offenders and hence sensors can be placed on these components to predict their future breakage.

3.3.2.3 Collect Asset Live Data

Now that the key components have been selected the task is to collect data for the DT from the physical asset. This information can be displayed on dashboards on HMIs or desktops or tablets on-site or remotely. The As-Is health of the key components can be seen by key stakeholders in real-time. This is a powerful tool; the digital copy of the asset creates new information and the asset is no longer a disconnected black box, but it is a cyber object, and more is known about its health than ever before.

Now that the components have been selected on key assets, sensors need to be placed on the asset. The sensors extract the data to the cloud. Wireless, glue-on sensors can cover a wide range of data types. They can measure tri-axial Vibration, Temperature, Amps, Speed and integrate with any other data like oil condition or process information.

The sensors selected for the manufacturing facility have a 5-year battery life and use ultra-long-range Bluetooth, with a range up to 200m. Once the data has pinged from the sensor to the gateway it can be uploaded to a PC or cloud server using Wi-Fi, ethernet or 4G. This robust and cost-effective sensor pack allows an asset to be digitized quickly and accurately as it pushes multiple data types to the cloud to build the DT.

The following section highlights some of the deep technical knowledge that is required to collect asset live data and prepare it for use in a DT. The detail shows the skillset required for this work and that engineers must become comfortable discussing this type of work if the I4.0 strategies are to be fulfilled.

3.3.2.3.1 Extract Transform, Load

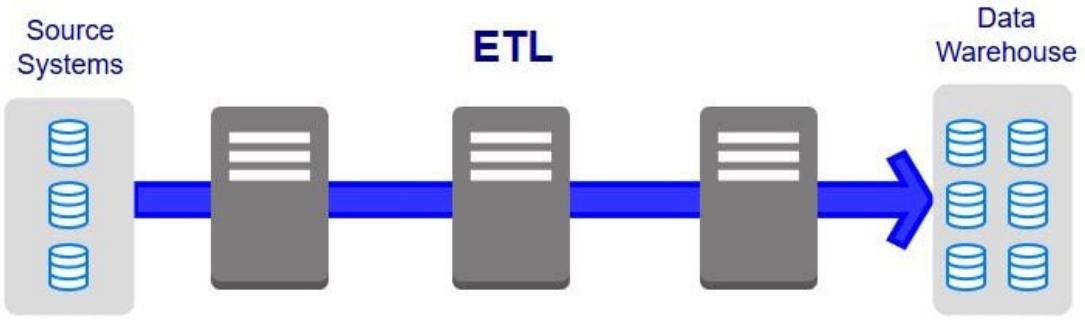


Figure 10 Extract, Transform and Load from source to storage [69]

Movement of data from a source area to a destination area and changing it into a compatible format is called Extract, Transform & Load (ETL). This is a key part of any I4.0 project as data is generally required to be moved from the source to a processing point to generate knowledge, this point can be at the edge or in the cloud.

These three steps can be broken down a bit more to describe the steps & decisions in each [70];

Extract	Transform	Load
Profiling	Mapping	Application Code
Access Layers	Lookup	Independent Code
Staging	Aggregation	Data Structure
Extraction System	Normalization	
	Business Approval	
	Transformation Location	

3.3.2.3.2 Extract

Before extraction of data from a source can occur profiling of the data should be carried out. It is the process of analyzing and checking the data to be moved is in the format as stated in documentation and the content is as agreed. This step can prevent issues further into a project when the source data may need to be cleaned or corrected as there is differences in the data to what the project team thought it would be. This work can lead to increased project costs and timeline delays.

Once the source data is known and as requested the data extraction can occur. This can occur in two formats; the source system can make a copy for extraction or an extraction system can get the data from the source. The source system will have another function so adding extraction work to this is not preferred option. Adding a dedicated extraction system with minimum impact onto the source system can work well.

Staging the data is storing copied extracted data on the source, intermediate or target server. This can be important for audit trails but can increase the cycle time.

If an extraction system is going into a source system to get data, it will need to move through various access layers protecting the data. It must be necessary for the system builder to know all the layers involved in this step and secure proper access through the security layers. The access layers can include; data

structure, application layer, operating system layer, server layer, firewall layer, network layer and cloud layer [71].

3.3.2.3.3 Transform

The transformation step changes the data from the source format to the desired destination format. This transformation must be agreed with the end customer to ensure the correct target format is selected.

This change can occur at the source point, i.e. ETL, or after the load step at the destination, i.e. ELT. This decision would be driven from a technical perspective to where is the best location to carry out this transformation work.

The transformation step is very much project specific as it can be a simple step, or it could involve a large collection of data from multiple sources requiring various changes.

Simple mapping is movement of data from point A to point B with minor adjustments. Lookup is sourcing a data point in a data field what must be searched for, found and copied out. Again, once located it may need minor or major adjustments before it is loaded to the destination point.

More detailed and time-consuming changes to the data may be required. For instance data aggregation and normalization involving retrieving multiple pieces of data for a single target point or vice versa.

3.3.2.3.4 Load

Loading of the copied data to the destination point can occur before or after the transformation. There may also be a staging point also at the load step as previously discussed.

There are two main methods to loading data; application code or independent code. Application code uses the destination application code that is in place to load the data into the target point. This would be the preferred option when possible as custom work should be avoided as it presents a new failure point for the system. The built-in application code, if it has the add-on, will allow the new loaded data to seamlessly fit into the application. Writing independent code may be required for legacy systems that cannot use an API.

When loading the data, it is important that the destination data structure rule set is kept on and followed to ensure data quality.

3.3.2.4 Generate Model

With sensors connected the next step is to use the data gathered to generate new knowledge about the asset. This step has a wide scope in terms of of what is possible, again it is important to reflect on the return on investment for the work being completed.

The predictive capability for the DT model could be as simple as a basic algorithm with min/max alarm parameters, an expert system or a system reliant on machine learning. It is important to highlight that only once data quality and ETL has been satisfactorily completed, should work on this step be started. There is frontloading of work required before this step can be completed. The business must be patient to allow data to build up before value can be seen from this step.

3.3.2.5 Train Model

Training the model is applicable to a DT that has machine learning capabilities. This may not be the case particularly for the first iteration of a DT created. However, as the experience of the team grows, and the data pool grows larger the introduction of a machine learning model will amplify the benefits of the DT.

Once the raw data has been extracted, transformed and loaded it is now called working data. This working data is checked for patterns as part of the machine learning process. The working data is split into training data and test data. The training data is used to train the model for the faults and issues that can be shown by the data collected. The model parameters then selected from the training set are tested on the test data to assess the machine learning models capabilities to predict issues.

3.3.2.6 Run Simulations

With a model trained and selected the next powerful step of the DT can be running simulations to investigate possible future states of the asset's health. With live data displayed the DT offers value, but with models and the potential to run simulations the DT can offer even more insight. Through classification fault issues can be categorized, regression can then be used to predict the probability of issues and simulations can be run on collected data to show future outcomes. The DT is now a potent tool for the maintenance group, showing knowledge and predictions never possible before. Academics in laboratory settings have started to simulate real-time production lines to create these models [72].

3.3.2.7 Prognostics

With models and simulations predicting probability levels of events, the ability to execute prognostic work is now possible with the DT. Combining this work with MTBF & MTTR the health management of the assets are now in a proactive state. With this knowledge available the importance is now placed on what decisions are made based on this knowledge. This decision-making process could be made by human-only, but a better solution is again to use the DT to the facilities advantage. The DT can include a decision support system, this aids the worker in making a more measured decision by using relevant data to inform their decision and remove any human bias that could negatively impact the decision process.

3.3.2.8 Failure Mode Effect & Critical Analysis

With a prognosis now given, there is a requirement to react to this information. The DT will contain a risk evaluation system based on the known failure types. The possible failures will be ranked on severity and probability of the occurrence based on the data. This system will then output a score for the possible failures.

3.3.2.9 Decision Support Systems

The score from the FMECA system will feed into a decision support system (DSS). The DSS unites human skill with digit processing power capabilities to provide efficient management of data and present data backed recommendations for the human [73]. The DSS will apply rules based on the

available event data and sensor data, and analytical and simulation software [74]. It will infer knowledge and recommendations about upcoming events. This informs the engineers of recommended actions. Benefits of the DSS include [75];

- It is an automated system and can continuously monitor the condition of a system. Whereas human experts may not be able to continually monitor a system
- The system can give an explanation as to how it arrived at a decision. It can give a reason why certain rules were selected
- It can also give a probability with all the decision options it gives with the likelihood of each being correct based on previous data.

This workflow goes through a person as the manufacturing industry will be slow to allow equipment to alter their state without prior approval from a person. Also, in relation to maintenance, most actions will require human interaction, such as replacing worn or broken parts.

The process of the DSS making a recommendation follows a set number of steps, these include [76];

- Data Source storage
- Data Warehouse modeling
- Analysis and review of extraction results
- Generation of reports and results
- Decision recommendation

Once a query is entered into the DSS the defect issue will be qualified to the relevant grouping. The cause of the defect will be identified, and ways or methods to eliminate it will be selected. From the selection choices the optimal solution will be chosen [76].

With a recommendation from the DSS system the worker can accept this response or if desired interrogate the decision by reviewing the data in the DT and possibly running further simulations. A Smart Maintenance Decision Support System is predictive tool that aims for “near-zero breakdown performance” as part of a CPS [77]. This goal is not currently achievable but smart systems with more data will mean ever increasing accuracy moving forward.

3.3.2.10 Recommendation to Engineer & Work Order Creation

With a recommendation from the DT, the engineer can decide to green-light the decision to create a work-order (WO) to carry out the remedial work. The DT should be linked up with operational and maintenance management system to create and assign the WO automatically. This approach fulfills the approach to create a CPS where the DT is interacting and feeding information to different systems in the enterprise.

3.3.2.11 Feedback

The WO task will be carried out by a maintenance technician. Once this work is complete it is important for the accuracy and improvement of the DT that event data from the works is inputted back into the DT. This information would, in an ideal state be filled in on the technician’s tablet. This data would get added to the data lake in the correct database (DB). With the data stored in the

correct DB for each asset, the DT can then pull on this feedback event data when making future decisions about the asset.

This feedback is vital for the DT as it will improve the DSS accuracy, highlight when a wrong decision has been made and event data is a pivotal source that can be combined with sensory data to create a clearer outline for the asset's health.

The event data can include a data thread which would allow the technician to state if the action taken was the correct one. It would allow feedback to say if the works occurred too soon, too late or at the correct time. It should also feedback if the works were the correct action or if different remedial work was required. This event data can then be used to retrain the models and hence improve the DSS accuracy for future decisions. The more the system is used the more accurate it should become.

If the DSS has recommended the wrong works, including the person in the loop adds another safety factor. A fully automated decision may change something in error. Including the person in the control loop allows the tacit knowledge of the worker to be added to the loop. If the wrong recommendation has come from the DSS, the event feedback data can state this and correct it for future problems. Over time the role of the person can be reduced as the DSS gets more usage and increases accuracy.

Event data is also a rich source of knowledge around an asset's health. To rely solely on sensory data is an oversight. Event data informs the DT of what happened and the maintenance work to fix it. It requires some manual entry, however each entry, if the data quality is high, reduces the likelihood of repeat issues.

3.3.2.12 Parts Ordering

Although not directly part of the DT, automatic parts ordering should be included in the DT framework. This part of the system sits within the larger manufacturing CPS that the asset DTs would reside in. Once the technician has completed their maintenance work, if a new spare part is required for stores the CPS should order it. The technician's tablet should show inventory levels and if after using the spare part another is required the software should automatically order one from the supplier.

Where the value from the DT would come is if the part being replaced has data on its lifecycle. This could be extracted from the DT and offer the maintenance team, supplier and vendor valuable insight into that parts health during its lifecycle. The data could be compared with similar parts in the facility but also cross industry to compare its performance. Valuable insights can be gained from this method, it may emerge the data shows the parts life was shorter than expected and under more stress than normal and that this should be investigated more to improve the overall asset's health that the part resided in. or it could show the part lasted longer than expected.

3.3.2.13 Digital Fleet

The DT for the asset is enriched overtime as more data is captured. With event data for all maintenance issues captured and added to the DT data set, the DT becomes stronger as a tool for the enterprise. With hundreds of assets in the facility and similar assets in global locations the DTs become a powerful analysis tool. The DTs for similar assets become a digital fleet, this fleet can be used to

compare unit performance across sites, forecast part replacement and failures based off fleet data and simulations, forecast global capacity in assets and review if capacity can be met. This benefit of the digital fleet is that it enables enterprise level decisions to be made based on data-led analysis helping to achieve better results.

A digital fleet will allow comparison of similar assets across sites. Historical and current data can be pulled up on any DT. The data will likely be refreshed every 15 minutes, so close to real-time, but not live. Live data would not be required though for this purpose. Engineers can compare performance of units over the previous year or since new parts have been placed in two different units. The data will show if one is performing better than another, this can start a conversation with the suppliers to improve the underperforming unit. Until now these types of conversations couldn't have happened as the visibility simply was not there.

With a dataset showing fleet performance, and event data showing when parts have been replaced. There will be the potential for accurate forecasting of when parts need to be replaced. This will replace inefficient time-based maintenance, that removes parts that are ok and worse replacing a part after it has broken. The analytics team will be able to run simulations based on the data for the fleet, these forecasts can run using different input parameters based on different scenarios and show various failure modes.

This digital fleet data could also be used in the wider enterprise CPS. The Global Supply Chain analytics team can simulate predicted future capacity based on proposed maintenance schedules. Global capacity for various assets will rise and fall as units go offline. It is important for the business to know the global capacity to make sure it can meet demand. The digital fleet can feed its data into this make-up.

3.3.3 Iterative Process of Framework & Modular Design

The design of the framework includes multiple steps, each section warrants an independent project in its own right. This structure was chosen on purpose, based on lessons learned from previous DT work that failed to be implemented due to its scale and complexity. Work on the DT should start at the first step by collecting data. Only once quality data has been collected can the remaining steps be completed. The quality of the data decides the quality of the DT, which decides the DT outputs.

Work can begin on different steps in a staggered manner, it is important that communication between the different work teams is kept open to ensure the overall aim is completed. This method of work creates value for the business sooner, take for example data collection. If the data has been collected by the facility, it can be shown on dashboards to engineers. Manual steps can be taken based on this new knowledge, the data can be imported to excel for manipulation easily. The alternative of waiting a longer time for the full DT to be completed has many pitfalls.

When work is being carried out in installing the parts of the DT, previous parts can be iteratively improved if required. This allows sections to be substituted for better technology as it becomes available. This plug and play swapping of parts allows the DT to be improved overtime. This continual

upgrade process is seen with all software systems nowadays. As bugs are discovered upgrades are carried out automatically.

This chapter discussed the previous smart manufacturing work executed in the facility, including the design of a DT and the lessons learned from the work. It highlighted some of the difficulties in constructing a DT and why there are not more of these assets in the facility, which were two of the main research questions for this thesis.

Issues arose across several areas that were categorized into business, human factor, project related and vendor specific that all hampered work and are likely reasons there are not more case studies of industrial DTs. Learnings directly from the DT work were that; data projects are more complex than the conceptual ideas first thought, they are expensive if the scope is not kept in check and the projects require domain knowledge.

As part of the active research model, the information gathered was reflected on and a plan was put in place for the next cycle of work. The response was to develop a DT with a bottom-up approach using the framework discussed in the next chapter. This bottom-up approach was in response to the lessons learned from the previous work where the large projects had not succeeded. This work following seeks to answer the third research question of the thesis which was “how would you build a maintenance digital twin?”

4 Case-study – Introduction of a Digital Twin

With a framework now in place on how a DT should be introduced the focus moved to the practical introduction of a DT to the facility for assistance to maintenance staff. A maintenance DT would help move the maintenance strategy from reactive to proactive. The DT will monitor key components of selected assets to enable usage-based maintenance strategy.

The manufacturing facility still has a body of work to bridge the gap between its current state and having a fully functioning DT with all capabilities. The facility has started work on the preliminary sections of the framework. The first step taken was to "Collect Asset Live Data". In an ideal state the DT will;

- A.) Provide condition-based monitoring and predictive capabilities
- B.) Provide simulations of possible future states of the asset to allow prescriptive capabilities

It will take a prolonged period to enable both A.) and B.) capabilities. Work was carried out to connect condition-based monitoring sensors to a selection of key assets in the facility as a pilot group of assets for testing.

CBM sensors have been added on a selection of assets to ascertain if the data connection can be successfully maintained and that the real-time data offers some findings. The sensors installed communicate via Bluetooth to a gateway on a Raspberry Pi, from here the data is transferred to the secure company cloud server. In this test phase only vibration data has been collected.

Fast Fourier Transformation of the collected vibration data occurs to convert the data to applicable graphs for analysis. Vibration data is the first data type selected for collection as it tells the most about an asset's health [44, p. 8], other data types such as temperature, current & oil analysis can be added in the future. These can then be combined with different data sets such as production and supply chain information to create a complete overview of the asset.

The analysis of this data is manual at this stage, only once a large block of data is built up can a data model be trained. This will allow the DT to take a bigger form where; data is fed into it, simulations are run, and prognostics performed automatically. The assets selected for the pilot case-study are high volume machines and equipment with recurring breakdown issues. The aim is to discover issues in their infancy and prevent major events.

This collection of real-time asset data is the foundation upon which all the analytics work of the DT can be completed. It is important also to pull event data into the portfolio. Event data describes any significant event that has occurred for the asset. It could range from commissioning, breakdowns and repairs. This data differs from sensor data as it will likely contain text and be a description of the event and the works carried out by the technician.

Andrew Jardine et al. raise the point that event data is just as important for a CBM strategy as condition monitoring data [46]. So, the plant is making a concerted effort to improve event data logging. This will help create the overall picture for when events do occur. The data collected on these assets will be checked against ISO guidelines for equivalently sized healthy machines.

4.1 DT Framework Usage

As part of the case-study a selection of key assets had sensors added to them to show the business the benefit the DTs could have to both the process

and the facilities group. 28 assets were selected after review by key stakeholders, the assets represented critical assets for the business across different equipment types.

The assets were mills, tumblers, grinders, air handlings units (AHU) and centrifuges. This mix allowed the group to show the benefit of sensing different assets and the knowledge that could be gained about them. No automated key component directory is in place yet, so SME knowledge was used to place the sensors on the correct components to be monitored. The sensors used were Erbessd Phantom V Tri-Axial Vibration sensors. The sensors were placed on the spindles, shafts and gearboxes as applicable.

Other types of sensors will be added to the DT as it evolves, vibration covers the most maintenance issues of the available sensors, this will be discussed more in following sections.

4.2 Predictive Maintenance Program

Predictive Maintenance strategy was discussed in section 2.3, the sensors added in this case-study begin the process of implementing predictive capabilities for critical assets. The benefits are listed below, and with emerging technology it is easier and easier to apply this program. The techniques and tools used in this strategy are discussed and certain parts are realized as part of this case-study. An interesting outcome is the almost instantaneous monetary value this work showed, highlighting the need to extend the program.

Paresh Girdhar stated “The advantages of predictive maintenance are accepted in industry today, because the tangible benefits in terms of early warnings about mechanical and structural problems in machinery are clear. The method is now seen as an essential detection and diagnosis tool that has a certain impact in reducing maintenance costs, operational vs repair downtime and inventory hold-up.” [78, p. 4]. Victor Wowk agreed with the above statement outlining three fundamental benefits to a Predictive Maintenance Program [44, p. 286];

- Only unhealthy machinery is stopped
- Machines can run past regular overhaul dates
- Early defect issues are found and can be fixed under warranty when they would have otherwise gone unchecked.

4.3 Digital Twin Data Sources

Modern instruments have greatly advanced our capability to acquire dynamic data. The interpretation of this data is the weakest link in the analysis chain. There is more information available than we can interpret [44, p. 2]. Hence it is important that an automated analysis system is setup such as a DT that helps workers make decisions about the physical asset.

As part of a Predictive Maintenance Program the correct sensors need to be mounted to create the DT and inform workers of key components health. Analysis can be gathered from many sources; the following is a list of some of the main ones;

- Vibration analysis
- Oil and debris analysis
- Thermography

- Ultra-sonics
- Temperature
- Current
- Voltage

All these data types cannot be added at once, so it is important to add the data source that adds the most value and covers the key issues first. Vibration data is the data source that does this, the table below shows the key maintenance issues and what data sources cover each. The table clearly illustrates why vibration data is the most important.

Table 6 Key maintenance issues and data types that highlight them [44, p. 8]

	Temperature	Pressure	Flow	Oil Analysis	Vibration
Unbalance					X
Misalignment	X				X
Bearings	X				X
Gears	X	X	X	X	X
Looseness				X	X
Noise					X
Cracking					X

Vibration data can show a whole array of maintenance issues that none of the other data types can. For this reason, vibration sensors were the first sensors mounted on the selected units.

Vibration analysis can be used for diagnostics on non-critical assets and with analytics can be used in a predictive format. Diagnostically the vibration data can be reviewed after an event to tell SMEs why the failure occurred. For predictive capabilities the vibration data can be monitored, and it can show degradation signs in components and inform SMEs of future issues that are likely to occur. This use of the tool optimizes capital spending on sensors.

4.4 Vibration Analysis

There are many types of sensing in a Predictive Maintenance Program like; Oil and debris analysis, Thermography, Ultra-sonics, Temperature, Current and Voltage. Most important however is Vibration Analysis, this subject is a full independent field of engineering that requires a body of knowledge itself. The main points of the subject which will be discussed are; Time-wave form analysis, Frequency domain analysis, High Frequency analysis, Velocity RMS and Phase Analysis.

Vibration analysis is an important tool that can be used to reduce or eliminate recurring machine problems [78, p. 10]. Vibration analysis is a tool

used to track issues that may be arising on equipment. It is used to prevent catastrophic failures and also to avert the following; downtime, higher cost of repair, lose production time, higher production costs, secondary damage & potential safety incidents. It can detect the earliest signs of issues and beginnings of component failure. The analysis can highlight issues on gears, bearings, shafts pumps. Mounting position and orientation are important and different issues can arise at different points on a piece of equipment. Hence it is important to place sensors at the correct location and in the correct orientation.

4.4.1 Vibration Introduction

In an industrial setting. vibration is the motion of a machine or its components away from and back to a set point [78, p. 13]. Each disturbance or motion is cyclical, and this can be studied in wave form. Multiple disturbance and components mean waveforms combine, this is also called superposition, this is shown in Figure 11 below.

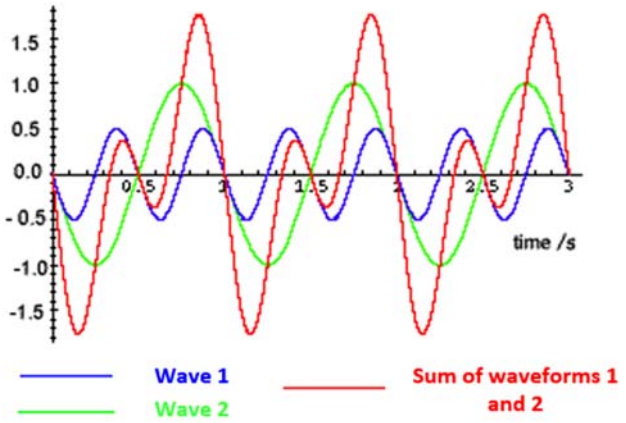


Figure 11 Superposition of waves [79]

Studying these waves is how vibration engineering findings are discovered. There are three vectors that can be used to study the waves; displacement, velocity & acceleration. A force can be examined using either of these three. For sine waves there is a 90°-sine difference between each [44, p. 40]. The application of converting a signal from displacement to velocity or velocity to acceleration equates to differentiation as a mathematical process. The inverse conversions also equate to mathematical integration. From a functional perspective when carrying these calculations out in the field using equipment it is easier and more accurate to integrate [80]. Hence accelerometers are the preferred tool of choice for vibration measurement.

When studying waveforms frequency, wavelength, amplitude and phase can all be analyzed as part of the work. Amplitude is a key measurement unit when comparing vibration waves. The amplitude is the first sign to show the health of the machine, generally the greater the vibration amplitude the increased risk of an issue [78, p. 21].

4.4.1.1 Displacement, velocity & acceleration

With a unit selected there is still the question of which vector to monitor, displacement, velocity or acceleration. The answer depends on the frequency of the signal you are looking at. Less than10Hz, displacement is favored as the amplitude for velocity and acceleration is too small to read. Generally, then from

10 into the hundreds, velocity is the easiest signal to analyze. Finally, greater than 1000Hz acceleration is used to check vibration signals [44, p. 42] [78, p. 21]. For most rotating machinery they operate between 10-1000Hz and most issues can be seen in this range [78, p. 22].

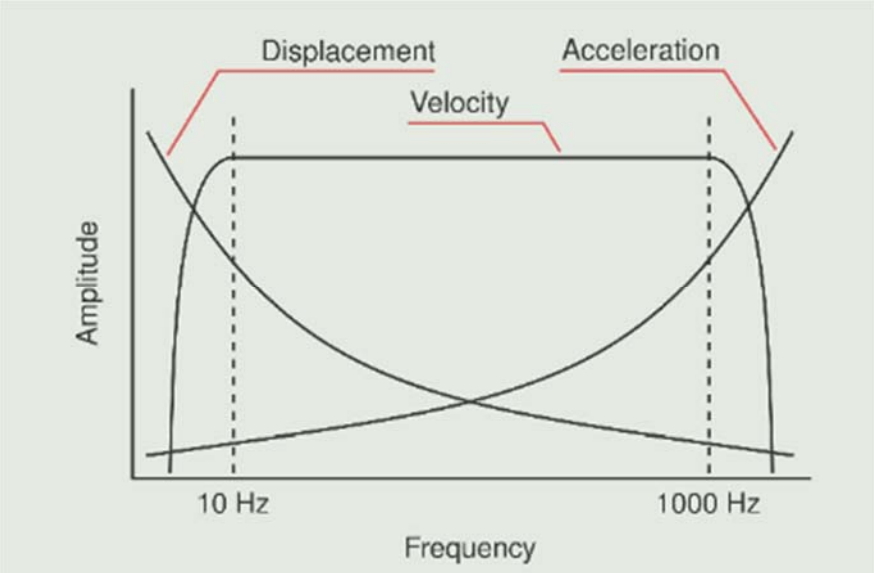


Figure 12 Vibration Amplitude as a function of frequency [81]

4.4.1.2 VelocityRMS

With most issues between 10 and 1000HZ velocity is the preferred measurement vector. When comparing vibration velocity, the average of the velocity cannot be used as this value may be zero in a balanced wave. Hence the Vibration Velocity RMS & Vibration Velocity Peak are used.

When vibration occurs to a mass it moves, the velocity of the mass changes. The velocity is at zero at the bottom and top ranges of motion. The velocity is at its peak as it passes through the neutral point on the axis. This maximum velocity is called the Vibration Velocity Peak [78, p. 20].

The root mean square (RMS) of a wave gives a magnitude for a wave where an average velocity value cannot, as explained above. The RMS amplitude format is valuable because it indicates the equivalent steady state energy value of an oscillating signal. Velocity RMS is the international standard for measurement of machinery vibration as set out by the International Standards Organization [78, p. 20].

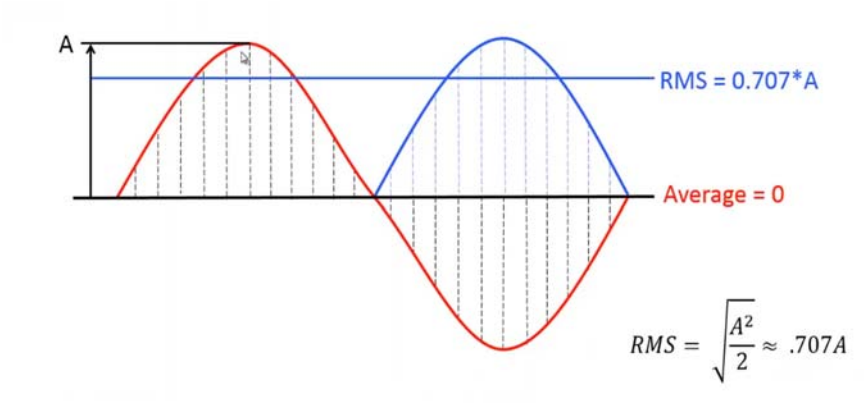


Figure 13 Waveform RMS Value [82]

4.4.2 Machinery Issues

Depending on the source the percentages vary slightly but it is clear imbalance and misalignment are the common rotating equipment issues. Paresh Girdhar states "80% of common rotating equipment problems are related to misalignment and unbalance" [78, p. 10]. Victor Wowk statistics vary slightly but the overall message is clear that imbalance and misalignment are the two main issues as shown in the table below. Vibration sensors can pick up these two mains issues and hence provide good coverage of the main faults. Adding the vibration sensor data to the maintenance DT will allow it to highlight issues that may have been missed by workers relying on scheduled checks only.

Table 7 Machinery Vibration Issues breakdown [44, p. 130]

Imbalance	40%
Misalignment	30%
Resonance	20%
Others	10%

4.4.3 Erbessd Phantom V Tri-Axial Vibration

To collect the vibration data sensors must be mounted on the physical twins. Shown below are details of the Erbessd Phantom V Tri-Axial Vibration sensor used. The sensor tracked vibration in all three axes X, Y, Z separately.

Phantom – V (Vibration)

Sensor Type	Accelerometer
Sensor Range	Triaxial: 8/16/32g; Biaxial: 20g
Sensor Accuracy	± 3dB
Triaxial Freq Range	10Hz–10 kHz (x,y), 5.1kHz (z)
Biaxial Freq Range	0.5Hz-15kHz (x,y)
Sampling Rate	Triaxial: 25600 S/s; Biaxial 44100 S/s
Floor Noise V122	0.12mg/√Hz (x,y), 0.2mg/√Hz (z)
Floor Noise V222	0.63mg/√Hz (x,y), 0.9mg/√Hz (z)
Internal Temperature Accuracy	±1°C (±1.8°F)
Operating Temperatures	-40 to 85°C (-40 to 185°F)
Storage Temperatures	-60 to 105°C (-76 to 221°F)
Size	47 x 33mm (1.85 x 1.3in)
Weight	100grams (3.5oz)
Battery Type	3.6V Li-TL-5935
Battery Life	2-4 years
Transmission Type	2.4 GHz BLE 5.0
Distance Range	150m, Line of Sight
Distance with Repeaters	Unlimited
Product Ratings	IP67 / MIL-STD-810G

Figure 14 Erbessd Phantom V Tri-Axial Vibration Sensor details

The sensor uses long range Bluetooth to transmit the data back to a gateway hub. From the gateway the data is sent to a database for storage using Wi-fi, ethernet or 4G. The sensors can be attached by adhesive or screwed in if required. They also have long lasting battery life. The intention is that Wi-Fi will be used to feed the data back to the databases. Currently the facilities Wi-Fi is being upgraded to allow this type of wireless data flow. In the early stages of the case-study work hand-held devices were used to collect the data for analysis as the Wi-Fi connection had not been authorized.



Figure 15 Erbessd Phantom Wireless Sensor Selection

The image below shows a sensor mounted and the three-axes required for full vibration maintenance insights. It is important to include all three-axes to

help identify what type of fault is present. Table 8 below shows all the assets and sensors used in the pilot phase of the case-study.

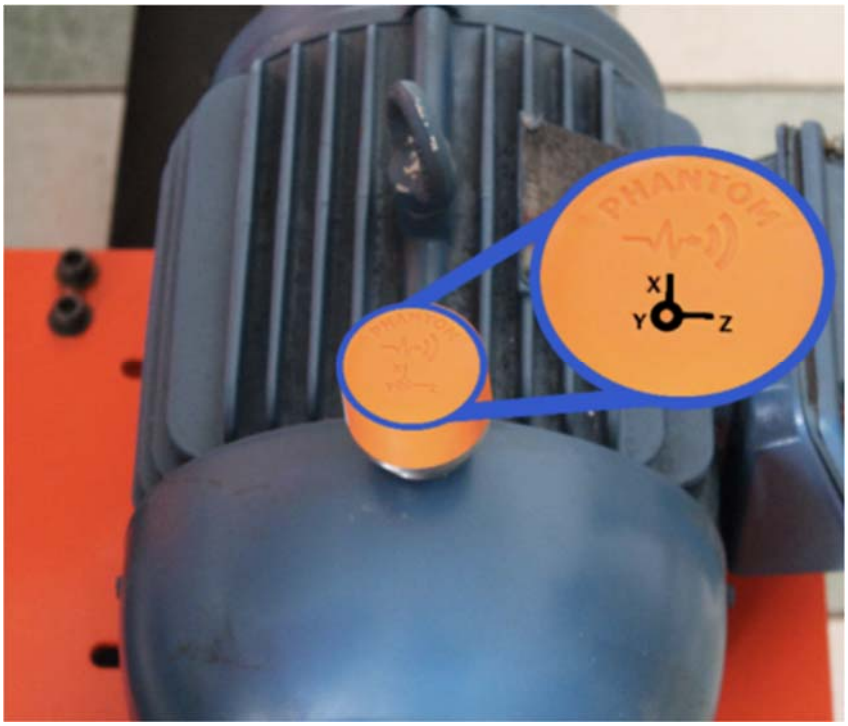


Figure 16 Erbessd sensor mounted with three-axes monitoring [83]

Table 8 List Assets & Sensors installed in the case-study pilot phase

Asset Tag Reference	Drive type	No. of Sensors	Sensor Serial #	Sensor name respective (Like drive end etc)	Erbessd Gateway #	Machine orientation	Sensor Orientation			Notes
							H	V	A	
DE826	Coupled	5	11 189 247 893	Spindle Mtr DEH	SN: 389245173	vertical	Y	X	Z	Flexible mount medium
			11 189 247 966	Spindle DEH			Z	Y	X	
			11 189 247 866	Xaxis H			Y	X	Z	
			11 189 247 821	Yaxis H			Y	X	Z	
			11 189 247 978	Zaxis H			Z	X	Y	
Mill0053	Coupled	5	11 189 247 802	Spindle Mtr DEH	SN:389245168	vertical	Y	X	Z	Flexible mount medium
			11 189 247 926	Spindle DEH			Y	X	Z	
			11 189 247 781	Xaxis H			Y	X	Z	
			11 189 247 921	Yaxis H			Z	X	Y	
			11 189 247 805	Zaxis H			Y	X	Z	
Tumb0015	Belt	5	11 189 247 975	Mtr DEH	SN: 389245168	Horizontal	Y	X	Z	Flexible mount medium
			11 189 247 865	GBX In H		Horizontal	Y	X	Z	
			11 189 247 795	GBX Out H		vertical	Y	X	Z	
			11 189 247 838	Shaft DEH		vertical	Y	X	Z	
			11 189 247 727	Shaft NDEH		vertical	Y	X	Z	
Tumb0016	Belt	5	11 189 247 999	Mtr DEH	SN: 389245168	Horizontal	Y	X	Z	Flexible mount medium
			11 189 247 933	GBX In H		Horizontal	Y	X	Z	
			11 189 247 943	GBX Out H		vertical	Y	X	Z	
			11 189 247 777	Shaft DEH		vertical	Y	X	Z	
			11 189 247 891	Shaft NDEH		vertical	Y	X	Z	
GRND 0062	Coupling	2	11 189 247 836	Spindle NDEH	SN: 389245168		Y	X	Z	Flexible mount medium
			11 189 247 799	Yaxis NDEA			X	Z	Y	
GRND 0058	Coupling	2	11 189 247 941	Spindle NDEH	SN: 389245168		Y	X	Z	Flexible mount medium
			11 189 247 804	Yaxis NDEA			X	Z	Y	
Centrifuge 1	Belt	1	11 189 247 806	Mtr DEH	SN: 389 245 171	vertical	Y	X	Z	Flexible mount medium
Centrifuge 1	Belt	1	11 189 247 830	Centrifige Brngs H		vertical	Y	X	Z	toothed belt
Centrifuge 2	Belt	1	11 189 247 824	Mtr DEH		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 475	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 578	Mtr DEH		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 471	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 473	Mtr DEH		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 737	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 743	Mtr DEH		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 497	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 782	Mtr DEH		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 969	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 831	Mtr DEH		vertical	Y	X	Z	
Centrifuge 1	Belt	1	11 189 247 819	Centrifige Brngs H		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 829	Mtr DEH		vertical	Y	X	Z	
Centrifuge 2	Belt	1	11 189 247 983	Centrifige Brngs H		vertical	Y	X	Z	
EFFL 0006	Belt	1	11 189 247 413	Mtr DEH		vertical	Y	X	Z	
EFFL 0006	Belt	1	11 189 247 519	Centrifige Brngs H		vertical	Y	X	Z	
EFFL 0007	Belt	1	11 189 247 649	Mtr DEH		vertical	Y	X	Z	
EFFL 0007	Belt	1	11 189 247 513	Centrifige Brngs H		vertical	Y	X	Z	
911 SFA	Belt	4	11 189 247 732	Mtr NDEV	SN: 389 245 174	Horizontal	X	Y	Z	Flexible mount,on springs, Medium
	Belt		11 189 247 895	Mtr DEH		Horizontal	Y	X	Z	
	Belt		11 189 247 998	Fan DEH		Horizontal	Y	X	Z	
	Belt		11 189 247 728	Fan NDEH		Horizontal	Y	X	Z	
911 SFB	Belt	4	11 189 247 967	Mtr NDEV		Horizontal	X	Y	Z	
	Belt		11 189 247 006	Mtr DEH		Horizontal	Y	X	Z	
	Belt		11 189 247 778	Fan DEH		Horizontal	Y	X	Z	
	Belt		11 189 247 886	Fan NDEH		Horizontal	Y	X	Z	

4.4.4 Time Waveform Analysis

The following sections discuss vibration analysis engineering. It is worth noting that these sections briefly cover each topic and there is a lot more detail to them as a subject if further detail is required. All these sections are only a small part of all the data sources that could be used in a DT. This highlights the broad spectrum of knowledge required to create a DT for maintenance. Issues that could arise from this wide knowledge base requirement will be discussed later in the document.

Time Waveform Analysis measures the sine waves from the moving parts in a motor, a single motion generates a single wave with a single frequency and amplitude. Additional components add more signals and is a more realistic reflection of an industrial motor. The full waveform is these combined waveforms, signals from each component are combined to give you the equipment wave form, separating and understanding each part of the wave form is key to vibration analysis. Figure 11 above shows this superposition phenomenon of combining waves.

Although time waveforms are hard to separate out and read, they do show up certain issues. The time waveform shows the continuous health of the unit, when reviewed over time it can show issues arising and changing the waveform. When performing vibration analysis, one must select the Fmax, the highest frequency on the spectrum and wave resolution. This will determine the time waveform seen, it is important that the correct values are selected so the time waveform offers insight into the unit health and is not at the incorrect scale.

Combining the time waveform and frequency spectrum analysis increases confidence in decisions made using vibration analysis. Measuring devices for time waveforms can be non-contact displacement probes (proximity probe), measuring the distance between the sensor tip and the shaft.

4.4.5 Frequency Domain Analysis

Frequency analysis, or spectrum analysis, separates out the different parts of the spectrum into separate wave components. Each component wave is at a different frequency. A Fast Fourier Transform (FFT) calculates out the different wave signals in a combined waveform. The power of the FFT spectrum analyzer in a machine monitoring environment is its capability to display vibration data in the frequency domain spectrum. This spectrum allows greater visibility to an asset's health [44, p. 17].

Analyzing growth in frequency peaks allows one to make a decision about maintenance issues. Changes in the spectrum over time allow you to track the machine's health. It is important not to continually compare frequency graphs but rather to see how the graph changes over an extended period of time, this shows you the health of the measured component and how that is changing with time.

An FFT Analyzer is a crucial instrument for digitizing vibration. Its importance is threefold [44, p. 90];

- Defects show up as specific peaks
- An overall number can be assigned to the graph, this allows easier comparison

- Small frequency issues sitting on larger waves can be split out and seen.

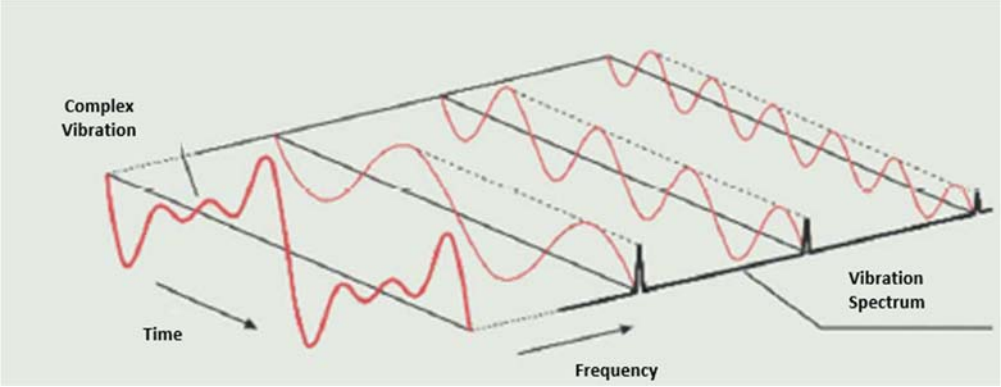


Figure 17 FFT processing of a complex vibration signal [81]

An FFT waterfall plot is a good visual manner to track the health of an asset over a prolonged period. Changes due to vibration along the full spectrum of hertz's can be seen.

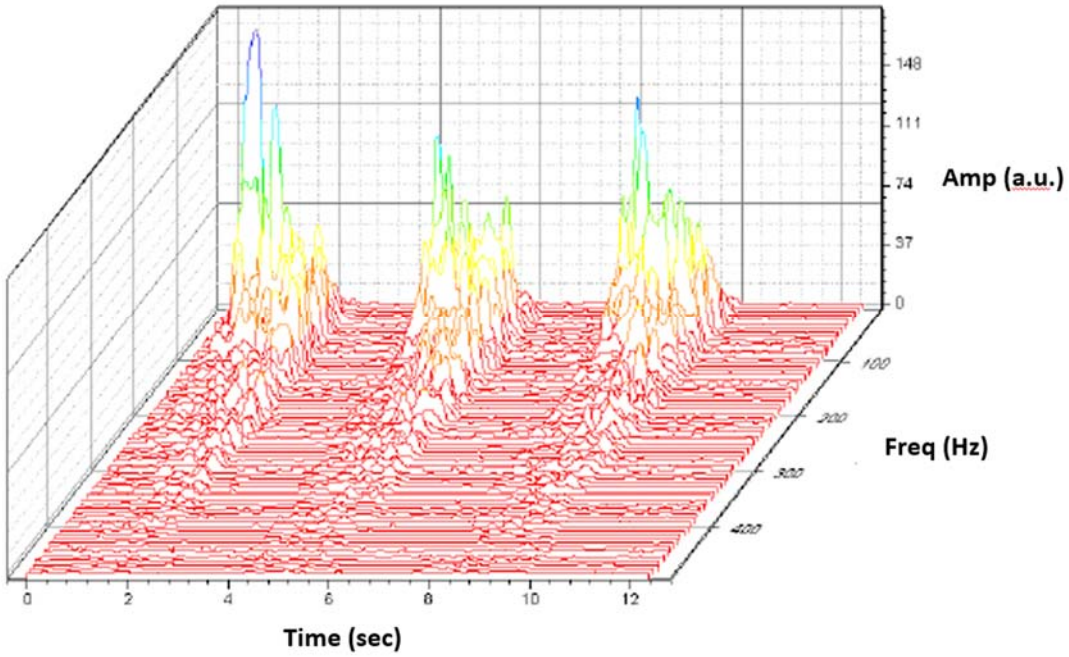


Figure 18 FFT Waterfall Plot [84]

4.4.6 High Frequency Analysis

High Frequency Analysis is analysis carried out at a frequency greater than 1000Hz (cycles/sec). High frequency analysis is an important investigative tool for certain types of issues. It works when the amplitude of the vibration is low, and the frequency is high.

This tends to occur with cracks, when the crack comes in contact with a ball bearing it causes a stress wave / shock pulse. This can be detected using high

frequency detection. Metal on metal clashes lead to high frequency vibrations that can be picked up with the correct examination of the data.

4.4.7 Phase Analysis

When there is an issue with a machine there is a motion in the machine, unique to each type of fault. That fault could be; unbalance, misalignment, bent shaft, eccentricity (orbit deviates from a circular path), looseness or another item. Spectrum analysis alone will not tell you the difference between all these issues. 1x 2x 3x peaks can be shown on FFT graphs from all the above reasons, so how do you pin-point the problem? Phase analysis must be used with the spectrum work to correlate to the exact fault type. It is possible that once an issue has been highlighted by the DT data that manual intervention will be required to confirm exactly what type of issue is occurring. Phase analysis would be part of this manual work.

4.4.8 Vibration Issues

The following section outlines the main types of issues vibration analysis can highlight.

4.4.8.1 Imbalance

As mentioned earlier imbalance has been highlighted as the primary concern with rotary machinery. Paresh Girdhar defined imbalance "as simply the unequal distribution of weight about a rotor's centerline. The ISO defines it as a condition that exists in a rotor when the vibratory force or motion is imparted to its bearings as a result of centrifugal forces. Correction of this unequal distribution of weight about a rotor is called balancing." [78, p. 135]. Balancing issues can occur due to manufacturing issues or overtime during operation of the equipment.

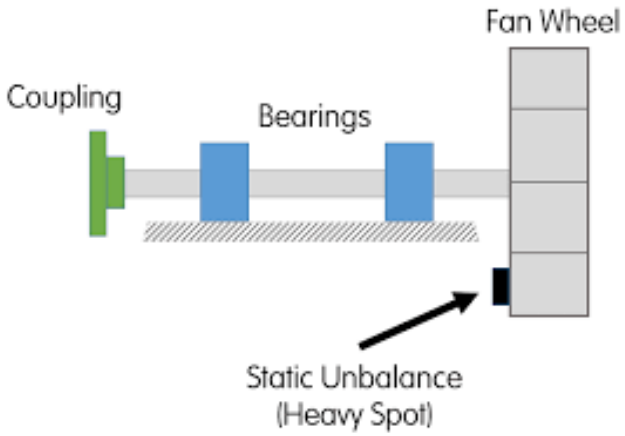


Figure 19 Shaft Imbalance [85]

Below is a list of some of the more prominent issues that cause imbalance;

- Lack of uniform density in machinery material
- Parallel boring issues
- Lack of symmetry in machinery shape
- Assembly errors
- Damage to machinery vanes or blades

- Distortion of material in the machinery due to temperature

The above concerns lead to an imbalance in the weight distribution for asset and over time this damages the system and must be tracked and fixed before it leads to catastrophic failures. The maintenance DT would track the components vibration data and if this increased over time because of a fault the DT would notify the maintenance team. Simulations could also be run on hypothetical imbalances to the component to see how and when the failures could be reached. This is the prognostics work that needs real-time data to be collected to feed into the model.

4.4.8.2 Misalignment

Misalignment is when the centerlines for two connected rotary-machineries are not lined-up. There must be collinearity between the coupled shaft axes [78, p. 45]. It is important that they remain parallel or elsewhere will commence leading to damage a component.

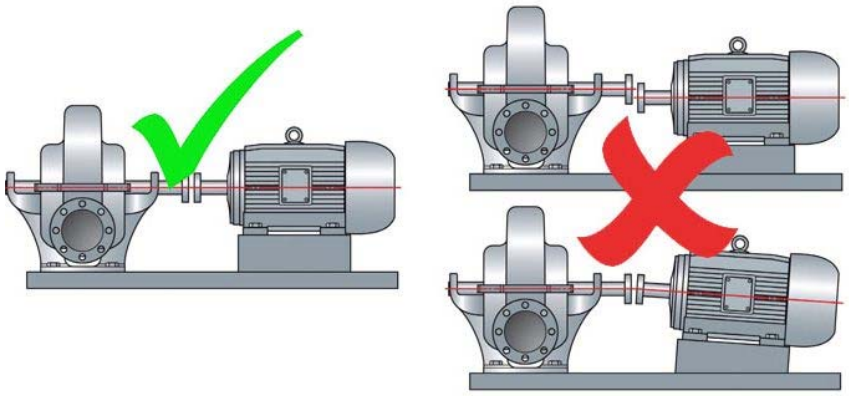


Figure 20 Shaft Misalignment [86]

During normal use alignment can drift due to following reasons [78, p. 145]:

- Thermal expansion of supports
- Stresses from pipes
- Deformation of structures
- Modifications in the sag of the rotor with a rise in temperature.

Misalignment can also occur internal to a machine with respect to bearings. The shaft movement needs to be concentric to prevent this from occurring. Again, like imbalance early detection of the issues for key assets is possible by creating a DT.

4.4.8.3 Resonance

Resonance is another common fault, less so than the previous two. Resonance occurs when the critical speed is reached and when the frequency vibration is equal to the natural frequency of an asset. Then the vibration amplitude rises significantly, much higher than expected compared to normal effects. These elevated-vibration amplitudes at critical speeds can be highly detrimental for any asset and must be prevented from occurring [78, p. 102].

4.4.9 ISO Standards

ISO-10816 allows comparison of machinery vibration data to assess the state of each asset. ISO-10816 outlines allowable vibrations and alarm or trip conditions for various machinery based on some statistical analysis of historical data gathered by ISO TC 108" [87]. Below is the table from ISO10816-1, vibration limits referring to general machinery broken into differing kW motor sizes.

Table 9 ISO10816-1 General Machines Velocity Range Limits

ISO10816-1 General Machines

Class I machines may be separate driver and driven, or could be coupled units comprising operating machinery up to approximately 15kW (approx. 20hp)
Class II machinery (electrical motors 15kw (20hp) to 75kW (100hp)), without special foundation, or rigidly mounted engines or machines up to 300kW (400hp) mounted on special foundations
Class III machines are large prime movers and other large machinery with large rotating assemblies mounted on rigid and heavy foundation which are reasonably stiff in the direction of vibration
Class IV includes large prime movers and other large machinery with large rotating assemblies mounted on foundations which are relatively soft in the direction of the measured vibration (i.e. turbine generators and gas turbines greater than 10MW (approx.. 13500hp) output

<table>
<tr><th colspan="2">Velocity Severity</th><th colspan="4">Velocity Range Limits and Machine Classes</th></tr>
<tr><th rowspan="2">mm/s RMS</th><th rowspan="2">in/s Peak</th><th rowspan="2">Small Machines Class I</th><th rowspan="2">Medium Machines Class II</th><th colspan="2">Large Machines</th></tr>
<tr><th>Rigid Supports Class III</th><th>Less Rigid Supports Class IV</th></tr>
<tr><td>0.28</td><td>0.02</td><td rowspan="3">Good</td><td rowspan="3">Good</td><td rowspan="4">Good</td><td rowspan="5">Good</td></tr>
<tr><td>0.45</td><td>0.03</td></tr>
<tr><td>0.71</td><td>0.04</td></tr>
<tr><td>1.12</td><td>0.06</td><td rowspan="2">Satisfactory</td></tr>
<tr><td>1.80</td><td>0.10</td><td rowspan="2">Satisfactory</td></tr>
<tr><td>2.80</td><td>0.16</td><td rowspan="2">Unsatisfactory (alert)</td><td rowspan="3">Satisfactory</td></tr>
<tr><td>4.50</td><td>0.25</td><td rowspan="2">Unsatisfactory (alert)</td><td rowspan="2">Satisfactory</td></tr>
<tr><td>7.10</td><td>0.40</td><td rowspan="5">Unacceptable (danger)</td><td rowspan="2">Unsatisfactory (alert)</td></tr>
<tr><td>11.20</td><td>0.62</td><td rowspan="4">Unacceptable (danger)</td><td rowspan="2">Unsatisfactory (alert)</td></tr>
<tr><td>18.00</td><td>1.00</td><td rowspan="3">Unacceptable (danger)</td></tr>
<tr><td>28.00</td><td>1.56</td><td rowspan="2">Unacceptable (danger)</td></tr>
<tr><td>45.00</td><td>2.51</td></tr>
</table>

This ISO tabulated data can be used to set preliminary alerts for maintenance workers. The DT can alert staff when an asset's vibration data is edging towards a limit. The ISO figures can be used initially and as the staff get more confident in their knowledge of the assets data these ranges can be

changed to be more specific to the asset. Using the data to increase the staff's knowledge is a good way to maximize the usage of components while being confident the likelihood of a failure is still low.

Based on Velocity RMS, values in mm/s, or Velocity Peak, in in/s, the guideline charts values to; good, satisfactory, unsatisfactory and unacceptable. Alerts and alarms can be added to real-time data collected to alert workers when the values are moving outside agreed parameter ranges.

4.4.10 Oil Analysis

Oil analysis is a useful tool when combined with vibration sensing. It can identify rotating equipment issues by studying the content of the oil itself. Abnormal wear elements that pass into the oil are identified by sampling the oil. This work is specific to each application; such as a compressor, boiler or CNC [78, p. 169]. This sensor could be added to an asset's DT and the oil content readings could be read remotely.

This is an example of adding data sources to create a better overall picture of the machinery's health. More data types like this mean the DT DSS will make better decisions as the DT will have more knowledge. To manually collect all this data is very labor intensive. Automatic data collection means a worker can remotely monitor the assets health and work with the DSS to make a maintenance decision for asset's anywhere in the world. The repair work recommended by the engineer can then be carried out locally.

4.4.11 Ultrasonic & Thermography Leak Detection

Ultrasonic leak detection uses high frequency, >20kHz, sound from pressurized gases leaking from equipment to detect the escape point. This sound is at a higher frequency than the range of human hearing. It can be used for leak detection and pump cavitation issues [88].

Another type of leak detection is infrared thermography detection. This uses a thermographic image showing heat energy emitting from an object. Small temperature differences can be detected using high resolution equipment [89].

Although these two tools would be carried out by hand, the event data from the field work could easily be added to the DT database. This creates a single repository for information relating to an asset and easy retrieval for future use by workers. it is important when integrating a DT into a maintenance strategy that all available tools are used for maintenance of the assets. Remote sensors alone may not be enough to alert staff to all issues.

This section introduced the DT framework and the predictive maintenance program and how the sensors for this program can help build the DTs. The following section discusses the initial case study discoveries and the benefit of initiating the predictive maintenance program, which is part of how to the overall strategy of building maintenance DTs.

4.5 Case-Study Discoveries

As outlined in Table 8 sensors were fixed onto a selection of machineries in the facility. The project has had mixed results thus far, some significant savings were captured by the install through the cost avoidance of preventing a change-out of multiple pieces of equipment, however the project has encountered some other obstacles. These issues are largely human factor issues that were discussed previously in the section on As-Is Assessment of a Manufacturing Facility maturity for Digital Twins. This section will cover the findings that created the initial monetary saving and the discuss the topics hampering further progress in the case-study work.

4.5.1 Case-study success

The first significant success of the sensor install project was by devices placed on centrifuges for process water pumps. Measurements were taken from the centrifuge motor drive end, motor non-drive end, centrifuge drive end bearings and the centrifuge non-drive end bearings.

4.5.1.1 Context of Data Collected

53 data points were collected from the 28 assets, this was because on some assets motor & gearbox data points were collected. Also data points for all 3 axes of X,Y,Z were also collected for certain assets. Data was collected throughout 2020, the data quality was impacted by the coronavirus pandemic as access to site was not possible at times, hence certain sensors were not tended to and went into sleep mode which meant gaps in the data stream.

The data was analysed by a vibration engineer specialist using a Erbessd MX 20 Digivibe hand-held analysis instrument as the facilities wireless monitoring platform for the DTs is not fully functional yet. Data is collected every 60 seconds and currently stored locally on the device. The intention is that the data will be sent to a platform in the cloud for automatic analysis, starting in 2021. This data then makes up one block in the construction of the DT.

Key data points include the asset name, sensor point / location, axis of measurement, timestamp, velocity, and acceleration of the vibration. Reviewing this data set allows diagnostic and predictive maintenance decisions be made. This is a small data pool, yet as seen from the use-case discussed still offered significant findings as shown below. This bottom-up approach is the preferred method of building towards a DT as the data and sensors added must pay for themselves as they are added rather than a large up-front capital expense on multiple data source and sensor types. The importance in the data collected for the business can be seen in the monetary savings and return on investment on the sensors added in the following section.

Most of the data collected for the year was in the satisfactory range limits and didn't warrant action. The data collected was compared against ISO-10816 standard tables, as mentioned in the previous section. It was Centrifuge 5's data in February that required further investigation and is discussed in the next section.

Table 10 Centrifuge Vibration readings taken 27/02/2020

Machine Code	Name	Point	Axis	Date	Vel Severity	Env Severity	Velocity mm/s	Acceleration Envelope gE	Acceleration g
Plastic Process Line B	Centrifuge1	1	H	27/02/2020 16:01:17	Yellow	Green	1.6434	0.001	0.186
Plastic Process Line B	Centrifuge1	1	V	27/02/2020 16:01:32	Yellow	Green	1.6499	0.002	0.253
Plastic Process Line B	Centrifuge1	2	V	27/02/2020 16:01:54	Yellow	Green	1.6843	0.004	0.263
Plastic Process Line B	Centrifuge1	3	H	27/02/2020 16:02:13	Yellow	Green	2.0077	0.006	0.126
Plastic Process Line B	Centrifuge1	3	V	27/02/2020 16:02:33	Yellow	Green	1.7416	0.005	0.188
Plastic Process Line B	Centrifuge1	4	H	27/02/2020 16:02:51	Yellow	Green	1.889	0.003	0.069
Plastic Process Line B	Centrifuge1	4	V	27/02/2020 16:03:06	Yellow	Green	1.7511	0.003	0.112
Plastic Process Line B	Centrifuge2	1	H	27/02/2020 16:04:34	Yellow	Green	1.9117	0.001	0.302
Plastic Process Line B	Centrifuge2	1	V	27/02/2020 16:04:56	Yellow	Green	2.0408	0.001	0.584
Plastic Process Line B	Centrifuge2	2	V	27/02/2020 16:05:33	Yellow	Green	2.2951	0.001	0.988
Plastic Process Line B	Centrifuge2	3	H	27/02/2020 16:05:53	Yellow	Green	2.7426	0.003	0.143
Plastic Process Line B	Centrifuge2	3	V	27/02/2020 16:06:11	Yellow	Green	2.3927	0.002	0.334
Plastic Process Line B	Centrifuge2	4	H	27/02/2020 16:06:33	Yellow	Green	2.4874	0.003	0.107
Plastic Process Line B	Centrifuge2	4	V	27/02/2020 16:06:52	Yellow	Green	2.1656	0.002	0.161
Ceramic Process Line B	Centrifuge3	1	H	27/02/2020 15:16:31	Green	Green	0.9123	0.001	0.318
Ceramic Process Line B	Centrifuge3	1	V	27/02/2020 15:16:55	Yellow	Green	1.1332	0.001	0.419
Ceramic Process Line B	Centrifuge3	2	H	27/02/2020 11:40:02	Green	Green	0.7538	0.108	0.185
Ceramic Process Line B	Centrifuge3	2	V	27/02/2020 15:17:18	Green	Green	0.7969	0.004	0.201
Ceramic Process Line B	Centrifuge3	3	H	27/02/2020 15:17:41	Green	Green	0.9791	0.003	0.079
Ceramic Process Line B	Centrifuge3	3	V	27/02/2020 15:17:57	Green	Green	0.8236	0.004	0.161
Ceramic Process Line B	Centrifuge3	4	H	27/02/2020 15:18:18	Green	Green	0.8805	0.006	0.077
Ceramic Process Line B	Centrifuge3	4	V	27/02/2020 15:18:34	Green	Green	0.7448	0.003	0.056
Ceramic Process Line B	Centrifuge4	1	H	27/02/2020 15:39:45	Green	Yellow	0.4871	0.158	0.242
Ceramic Process Line B	Centrifuge4	1	V	27/02/2020 15:40:01	Yellow	Yellow	1.2151	0.153	0.239
Ceramic Process Line B	Centrifuge4	2	V	27/02/2020 15:40:28	Green	Yellow	0.8812	0.238	0.408
Ceramic Process Line B	Centrifuge4	3	H	27/02/2020 15:41:19	Green	Green	0.5889	0.117	0.193
Ceramic Process Line B	Centrifuge4	3	V	27/02/2020 15:42:02	Green	Yellow	0.7575	0.272	0.468
Ceramic Process Line B	Centrifuge4	4	H	27/02/2020 15:42:40	Green	Green	0.5256	0.001	0.222
Ceramic Process Line B	Centrifuge4	4	V	27/02/2020 15:42:56	Green	Green	0.4569	0.001	0.171
Ceramic Process Line A	Centrifuge5	1	H	27/02/2020 15:11:49	Yellow	Green	1.3725	0.107	0.21
Ceramic Process Line A	Centrifuge5	1	V	27/02/2020 15:12:03	Yellow	Green	1.4349	0.117	0.205
Ceramic Process Line A	Centrifuge5	2	V	27/02/2020 15:12:29	Yellow	Yellow	1.4697	0.243	0.386
Ceramic Process Line A	Centrifuge5	3	H	27/02/2020 15:13:17	Yellow	Green	1.7037	0.006	0.182
Ceramic Process Line A	Centrifuge5	3	V	27/02/2020 15:13:43	Yellow	Green	1.6447	0.004	0.191
Ceramic Process Line A	Centrifuge5	4	H	27/02/2020 15:14:05	Yellow	Green	1.4278	0.005	0.153
Ceramic Process Line A	Centrifuge5	4	V	27/02/2020 15:14:22	Yellow	Green	1.4717	0.002	0.204
Ceramic Process Line A	Centrifuge6	1	A	21/11/2019 15:41:19	Green	Green	0.617	0.094	0.175
Ceramic Process Line A	Centrifuge6	1	H	27/02/2020 15:23:43	Green	Green	0.8048	0.001	0.081
Ceramic Process Line A	Centrifuge6	1	V	27/02/2020 15:24:00	Green	Green	1.0903	0.001	0.126
Ceramic Process Line A	Centrifuge6	2	H	27/02/2020 10:50:31	Green	Green	0.726	0.124	0.209
Ceramic Process Line A	Centrifuge6	2	V	27/02/2020 15:24:27	Green	Green	0.9277	0.104	0.171
Ceramic Process Line A	Centrifuge6	3	A	21/11/2019 15:43:07	Green	Green	0.3464	0.076	0.134
Ceramic Process Line A	Centrifuge6	3	H	27/02/2020 15:24:47	Green	Green	1.0204	0.048	0.082
Ceramic Process Line A	Centrifuge6	3	V	27/02/2020 15:25:06	Green	Green	0.9412	0.071	0.116
Ceramic Process Line A	Centrifuge6	4	H	27/02/2020 15:25:24	Green	Green	0.9774	0.024	0.045
Ceramic Process Line A	Centrifuge6	4	V	27/02/2020 15:25:48	Green	Green	0.8654	0.047	0.09
Plastic Process Line A	Centrifuge7	1	H	27/02/2020 15:31:30	Green	Green	1.0133	0.073	0.137
Plastic Process Line A	Centrifuge7	1	V	27/02/2020 15:32:04	Yellow	Green	1.8166	0.106	0.227
Plastic Process Line A	Centrifuge7	2	H	27/02/2020 10:44:43	Green	Green	1.0945	0.007	0.266
Plastic Process Line A	Centrifuge7	2	V	27/02/2020 15:32:31	Green	Green	1.0808	0.131	0.229
Plastic Process Line A	Centrifuge7	3	H	27/02/2020 15:32:59	Yellow	Green	1.2098	0.061	0.096
Plastic Process Line A	Centrifuge7	3	V	27/02/2020 15:33:15	Green	Green	1.0667	0.091	0.137
Plastic Process Line A	Centrifuge7	4	H	27/02/2020 15:33:40	Green	Green	1.102	0.04	0.069
Plastic Process Line A	Centrifuge7	4	V	27/02/2020 15:34:08	Green	Green	0.9618	0.05	0.083
Plastic Process Line A	Centrifuge8	1	A	21/11/2019 15:36:09	Green	Green	0.392	0.13	0.222
Plastic Process Line A	Centrifuge8	1	H	27/02/2020 15:07:12	Green	Green	0.8575	0.041	0.077
Plastic Process Line A	Centrifuge8	1	V	27/02/2020 15:07:26	Yellow	Green	1.2591	0.051	0.106
Plastic Process Line A	Centrifuge8	2	V	27/02/2020 15:07:53	Green	Yellow	0.9323	0.181	0.322
Plastic Process Line A	Centrifuge8	3	A	21/11/2019 15:37:37	Green	Green	0.2453	0.131	0.201
Plastic Process Line A	Centrifuge8	3	H	27/02/2020 15:08:18	Yellow	Green	1.1588	0.095	0.159
Plastic Process Line A	Centrifuge8	3	V	27/02/2020 15:08:39	Green	Green	1.0103	0.129	0.199
Plastic Process Line A	Centrifuge8	4	H	27/02/2020 15:09:09	Green	Green	0.9555	0.062	0.09
Plastic Process Line A	Centrifuge8	4	V	27/02/2020 15:09:31	Green	Green	0.9317	0.069	0.119

The table shows the readings taken are within recommended levels for the assets. The studied data led to the conclusion that none of the assets warranted changing out.

The drive belts were inspected as part of the survey. Some belts were loose, worn or in poor condition, these belts were tensioned, and Centrifuge 5 had its belt replaced. This issue with Centrifuge 5 was consistent with low amplitude peaks visible in its spectra as shown below.

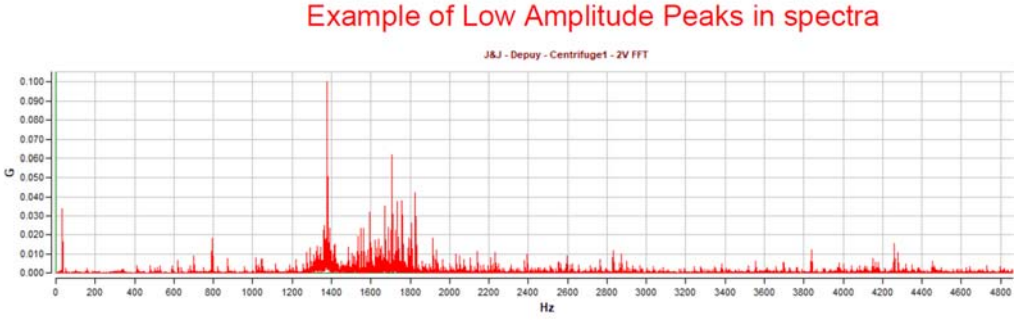


Figure 21 FFT spectrum for Centrifuge 5

Table 11 Centrifuge 5 readings before and after belt replacement

Vel Sev	Env Sev	Name	Point	Axis	Date	Velocity	Accelerat	Acceler
Yellow	Green	Centrifuge	1	H	12/02/2019 14:31	1.9889	0.123	0.208
Yellow	Yellow	Centrifuge	1	V	12/02/2019 14:31	2.2332	0.167	0.272
Yellow	Orange	Centrifuge	2	V	12/02/2019 14:32	2.1578	0.3	0.608
Yellow	Green	Centrifuge	3	H	12/02/2019 14:32	2.758	0.09	0.172
Yellow	Green	Centrifuge	3	V	12/02/2019 14:34	2.4304	0.084	0.155
Yellow	Green	Centrifuge	4	H	12/02/2019 14:35	2.3684	0.061	0.119
Yellow	Green	Centrifuge	4	V	12/02/2019 14:35	2.4197	0.096	0.155
Yellow	Green	Centrifuge	1	H	27/02/2020 15:11:	1.3725	0.107	0.21
Yellow	Green	Centrifuge	1	V	27/02/2020 15:12:	1.4349	0.117	0.205
Yellow	Yellow	Centrifuge	2	V	27/02/2020 15:12:	1.4697	0.243	0.386
Yellow	Green	Centrifuge	3	H	27/02/2020 15:13:	1.7037	0.006	0.182
Yellow	Green	Centrifuge	3	V	27/02/2020 15:13:	1.6447	0.004	0.191
Yellow	Green	Centrifuge	4	H	27/02/2020 15:14:	1.4278	0.005	0.153
Yellow	Green	Centrifuge	4	V	27/02/2020 15:14:	1.4717	0.002	0.204

The table above shows the reduction in vibration once the belt was replaced in Centrifuge 5. Velocity RMS values on the 12th of February range from 1.9889mm/s to 2.4304mm/s, these readings are from before the remedial works were done. After the works were completed the readings from the 27th of February range from 1.3725mm/s to 1.7037mm.s. As expected, the repair works have reduced the machines vibration level and improved its health and longevity.

Complete change out of these centrifuges was planned due to concern around the unit's noise levels. The approach for swap out of the asset's was based on tacit knowledge of the maintenance team. The data gathered from the sensors contradicted the tacit knowledge. The decision was made to retain the assets and swap out the belts as required. This is an early example in the DT framework of using a data-led approach to maintenance strategy decision-making. The data created an informed decision based on evidence of the asset's health and showed that even humans with expert domain knowledge are still susceptible to making mistakes.

4.5.2 Cost Benefit Analysis for Centrifuges

The value created by adding sensors to machines to create a DT for it will not be feasible if the finances of the works aren't viable. The early work carried out in this case-study show promising financial returns. The work so far has reduced business expense on purchasing new pumps, while the capital investment has been small and the potential savings around prevention for downtime cost to the business is high also.

Looking at the centrifuge pump example taking €75/hour as a labor rate, replacement of the pumps costs would be in the range of €70,000 to €200,000. This depends on the vendor who replaces the pumps, a local vendor would be cheaper than the (OEM).

Table 12 Option 1. Total Cost for OEM Pump Replacement

	Cost	Quantity	Total
Labor	€75	24	€1,800
Cost / pump replacement	€20,000	10	€200,000
			€201,800

Table 13 Option 2. Total Cost for Local Vendor Pump Replacement

	Cost	Quantity	Total
Labor	€75	24	€1,800
Cost / pump replacement	€7,000	10	€70,000
			€71,800

The manufacturing facilities maintenance team intended to replace all centrifuge pumps at a cost of between €70,000 - €200,000 because of suspected issues due to noise concerns.

Using the knowledge gained from the sensors added during the case-study work, this replacement of assets did not take place. The total cost to arrive at the conclusion the units did not need replacement, and completion of minor maintenance works instead was €4,450.

Table 14 Cost to Identify Issue & Complete Repairs to Centrifuges

	Cost	Quantity	Total
Install of sensors	€1,000	1	€1,000
Labor	€75	40	€3,000
Replacement of faulty belts	€45	10	€450
			€4,450

This saving shows the sensors have paid for their investment already. The saving achieved by the business is considerable and shows the benefit of digitizing equipment. The sensors now continue to collect data and can offer continued benefit to the physical asset. The return on investment is excellent and is a success for the project that will help with management buy-in as more capital will be needed in the future to fit out the DT fully.

The savings highlighted above do not include the added cost due to production loss due to downtime. This total cost of failure can far exceed the cost due to replacing a component. An example from the IIC Journal of Innovation highlights a failure in a production line costing €250,000 for the direct components but the full bill due to lost production downtime was over 50 operators over 5 shifts not working, with a total cost of failure in the region of €10 million [43]. While not all breakdowns will be this cost severe having a DT can offer value from a ROI perspective.

4.5.3 Summary of Case-Study so far

The case-study work so far has showed the benefits of trying to create DTs of assets while the process has still been hampered by issues. The cost of the DT work so far has been paid back by the savings highlighted in the centrifuge's repairs. The data from the sensors has not been sent to the cloud for storage and analysis yet. Fixing this matter has taken far longer than forecasted. The work so

far has also shown the importance of keeping the human in the feedback loop for the DT, as the worker must carry out the repairs highlighted by the DT.

The previous section highlighted the cost benefit of undertaking this work. The overall expenditure so far in the DT framework is low. No ML models, decision support systems or automated FMECA have been attempted. These are all next steps, up to now the expenditure has been on installing the sensors to collect data. The sensors selected offer a robust data source at good value and are a key component for a successful DT. If the data collection is too expensive or if the data quality is inconsistent due to sensor issues the DT as a tool will suffer. Cheap, reliable sensors have been sourced and used to collect the data upon which the DT can be built. This part of the case-study has been a success so far.

4.5.3.1 Case-study difficulties

Being in a regulated industry and using new technology in manufacturing facilities were raised as obstacles to the success of previous IOT projects in the facility. These two issues have arisen again as part of this project. The framework designed for the DT implementation doesn't have resolutions to these topics. Medical devices is a highly regulated industry, and this applies to all aspects of the shop-floor. The IT/OT departments until now have not had to connect and manage hundreds of edge devices that are pushing more data into the cloud. The work in the facility to connect the sensors to the cloud for data storage has been delayed by administration issues. So remote analysis of the data has not been possible, analysis has been done in-person. This is not the long-term solution, however adaptability in the face of problems is important and it has been vital to be able to highlight project successes using this data collected manually. The intention is to roll out a full web-based cloud platform, but this will take more time.

4.5.4 Future hybrid workers

The case-study work so far has shown the importance of having capable workers in the facility who can understand the output of the DT and act on it. Maintenance work will always require an amount of manual labor. The idea of lights-out manufacturing can be exaggerated today, this case-study has shown that information can be created by a DT. This can then be presented to a worker; they then must have the experience to act on the information. The hybrid engineers discussed previously will be well placed to be able to understand the process data while also being able to extract the data for the DT. A FMECA and DSS will help the worker make their mind up about a decision but to carry out the maintenance work the human will still need a level of expertise about the asset.

4.6 Case-Study Next Steps

The case-study work so far has proven its benefit to the business from a cost perspective. The method of completing smaller sections of work has meant that delays have not been as impactful if the project was bigger. The case-study has many areas to move towards. When reviewing the work completed and future direction of work, using the DT framework as a guide creates a clear view of the iterative growth of the DTs capabilities within the maintenance group.

In this section the development of a maintenance DT will be discussed by focusing on three iterations of the DT framework. Version 1 is the current completed work, version 2 includes more data sources, and the inclusion of the FMECA library and DSS, then version 3 adds in improved data analytics capabilities and connection to other applications used by the maintenance team.

Each section of the framework is a different tool that requires various skillsets to develop. The analytics team will focus on generating, training and running the DT models. Operational technicians will install the sensors that will provide the data and another specialist team will develop the FMECA analysis tool and the DSS. What is imperative is that as these tools are designed at different points in time that their connection to one another is highlighted by the project management team. By separating out the DT into smaller projects there is a danger that different teams working on tools within the DT will work in a silo. They will design in an isolated manner that could harm the overall performance of the DT. Hence it is the project managers job to thread the various tools together to create a seamless virtual tool.

4.6.1 Digital Twin Framework for Maintenance Version 1

The figure below shows the development work done so far in the DT framework. The green boxes highlight the work done for this version.

A pilot list of critical assets was selected to test the sensor capabilities. As previously discussed, this list contained a variety of assets to investigate how the sensors performed. Future versions of the framework will include a more thorough selection process based on a criticality screening process. Tacit knowledge was used to select the location of the sensors on the assets. An external vendor worked with the maintenance and process teams to review previous failures and high-risk components to decide sensor positions. The sensors were switched on and the data collection process started.

For version 1, only vibration data was collected. As previously discussed, this data type provides the most value from a maintenance perspective. The web-based data storage and analysis portal was not ready for use, handheld data readers were used to collect and analyze the data. This process will be replaced by an automated data collection and analysis tool. The iterative design process allows decisions to be made on data already collected rather than waiting for the finished solution to be ready. Investigation of the data graphs was done manually and checked against relevant ISO standards. Evaluation of the assets was carried out based on the data. Changes were made to the centrifuges as previously discussed.

The advanced detection of possible faults even using manually collected data is a powerful tool in prognostics work. The external vendor made recommendations to the maintenance team for remedial work based on the knowledge gained from the data. This recommendation comes from expertise knowledge and the BINDT Vibration Guide. This guide can in-time be integrated into a DSS tool. With a first cycle of the process complete the manufacturing facility maintenance team can begin to see the benefit of moving from time-based strategy to condition based maintenance.

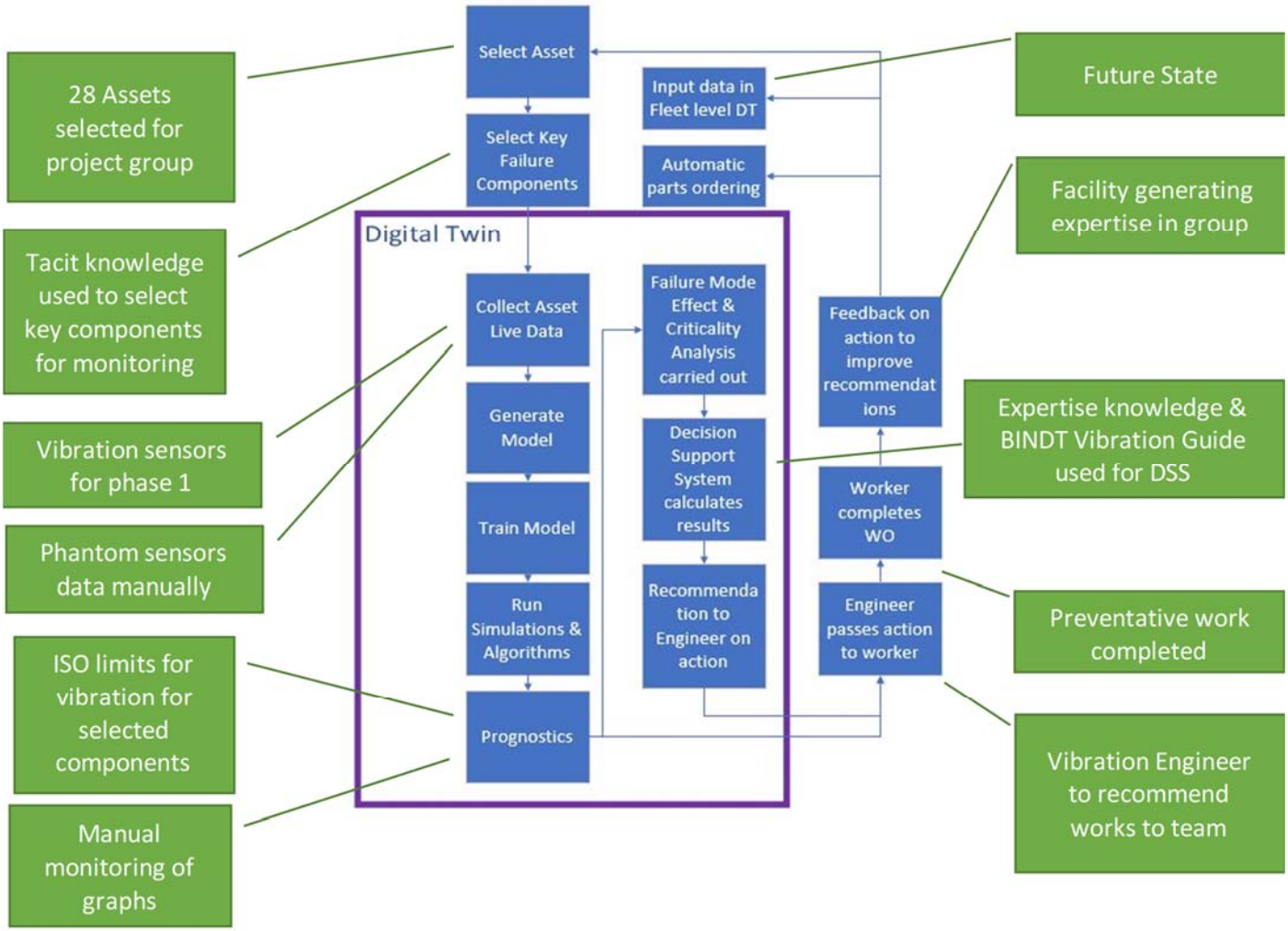


Figure 22 Digital Twin Framework Version 1

4.6.2 Digital Twin Framework for Maintenance Version 2

The DT framework below in Figure 24 shows the next steps to upgrade the DT. Projects to upgrade the tools can run in parallel and the existing DT can still function as parts are being added to it.

4.6.2.1 Selecting key components

Following the pilot phase, the maintenance team should draw up a list of assets which should have DTs created for them. This work can be linked to the current asset criticality assessment done yearly by the process and maintenance team. As part of this discussion it should be agreed the assets that would benefit the most from having a twin. By attaching the DT asset selection process to this asset criticality assessment key stakeholders will be allied and aware of the project work.

With the key components for the maintenance DT selected the sensors will be added to the physical twin. Future state DTs will draw on more data sources than just vibration, for version 2 these may include other scientific properties such as temperate, current, speed and oil properties.

4.6.2.2 Cloud platform for digital twin

Version 2 will have the data being collected in a cloud-based database. From here the team will be able to integrate the data remotely to perform prognostics. Alerts based on ISO standards and team preferences can be placed on the data graphs. The amount of data being collected will need to grow considerably before the DT can self-control.

With this version, FMECA and DSS capabilities will be added to the DT. The potential failure modes will be highlighted by the FMECA tool. This will then rank the possible failures. The DSS system, like an expert system will recommend a decision for the engineer. The system will operate like a decision support tree; “if this happens…then do this”. This system will be designed by the maintenance team and can pull data like production capacity and demand, spare parts inventory levels etc. All this data will allow the DT to make a more informed recommendation to the engineer. The engineer will have the final say as the asset cannot fix itself, however the DT is increasing the likelihood of making the correct decision for the asset’s health.

4.6.2.3 FMECA Tool

As part of the next steps in the case-study work, with key assets selected an analysis of component failure modes should be done. This work would create a Failure Mode Effect and Criticality Assessment library and a Failure Mode Analysis tool. The FMECA library would list all the failures types for the key assets selected to have DTs. Firstly, the FMECA tool would highlight the criticality of different failures and based on this, expertise knowledge and previous maintenance problems for the selected asset the team could decide what components they want to place sensors on. The Failure Mode Analysis tool would then highlight to the engineers the correct sensors to be added to the component. Different sensors track different issues and the Failure Mode Analysis tool shows which sensor tracks each type of fault. The figure below shows the sensor selection process for a pump.

The above work is done before the DT is created then the relevant data is collected to create the DT. Once the DT is created the FMECA tool can be used in real-time to score and rank potential issues based on the data being collected. Data could be fed into the tool manually or automatically based on the DTs development. The tool would be used on historical data to rate a possible fault. It could also be used on simulated data findings once this feature has been added to the DT.

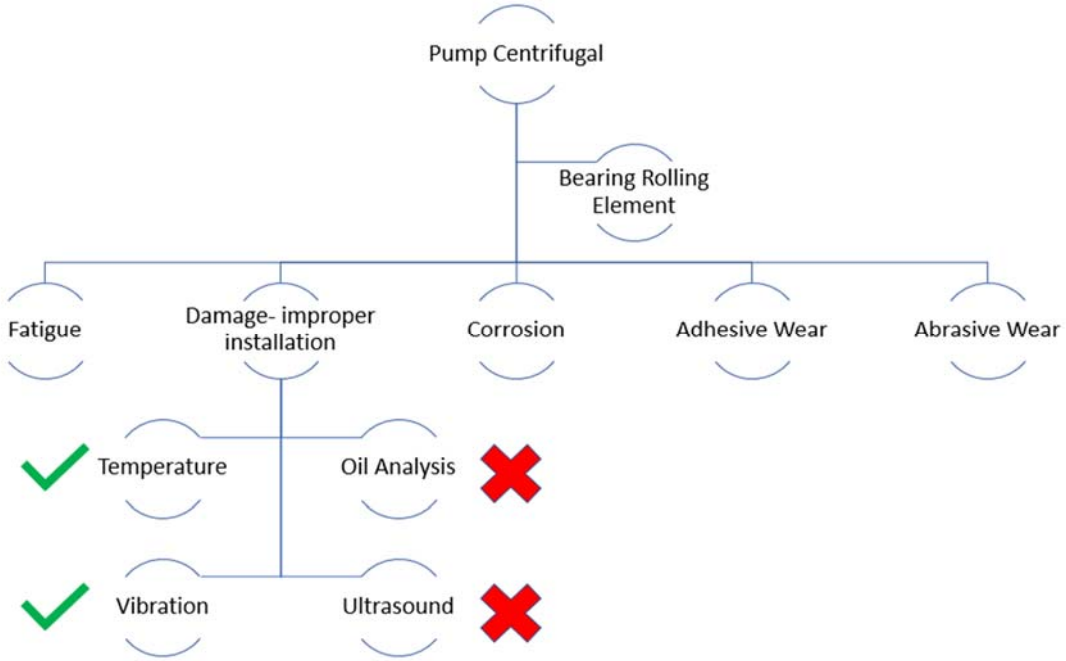


Figure 23 Example: Sensor Selection for damaged bearing in a pump using Failure Mode Analysis Tool

4.6.2.4 Schad Application integration with Digital Twin

The maintenance DT as it develops will be integrated with other applications used by the maintenance team in the manufacturing facility. One such application being tested now is Schad. Schad is an application that offers mobile access to asset maintenance data. The application has four main functions;

- Mobile access to work-orders
- Spare parts management
- Alarms & alerts for assets
- QR Code link to asset data

This app allows maintenance technicians mobile tablet access to data that previously required the worker to return to their desktop to complete the work. Now the technician can stay next to an asset and have the data at their fingertips. This mobile access to data combined with the DT creates a powerful CPS for the facility.

4.6.2.5 Mobile access

Version 2 of the DT framework would allow recommendations from the DT's DSS to be sent to an engineer remotely. The engineer can review this and assign

the WO to a technician as required. This mobile workflow is the future of connected factory workers and is enabled by the DT.

Once the maintenance technician has fixed the issue, they can use Schad to close the information loop. The technician can give feedback on the recommendation from the DSS of its relevancy. This improves the DSS as feedback improves its accuracy. Schad can be used by the technician if they used spare parts to replace a component in the asset. The spare parts management function can be used to automatically update spare parts inventory levels. This system can be linked up with a supplier and parts can be automatically ordered if levels are low.

4.6 Case-Study Next Step

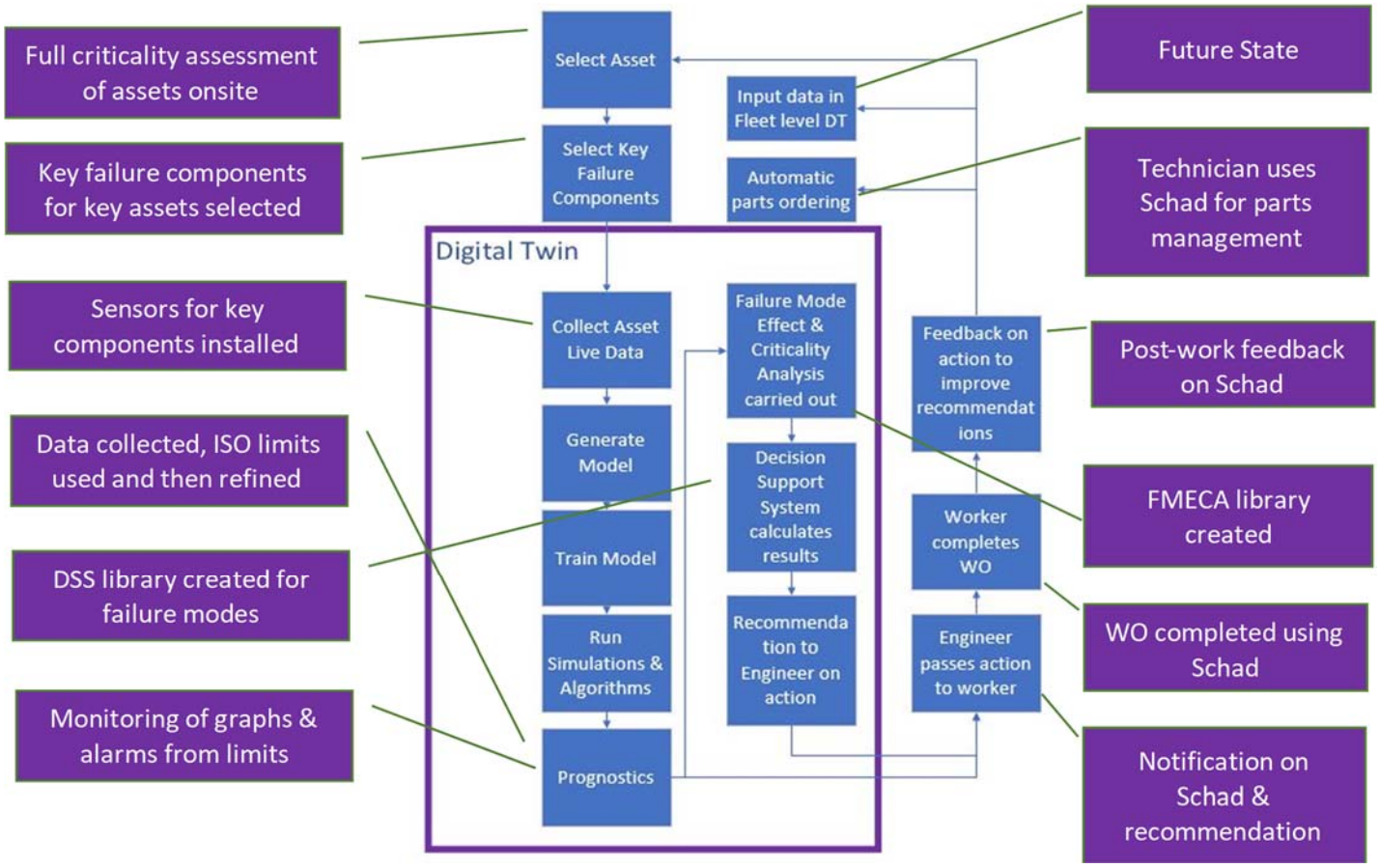


Figure 24 Digital Twin Framework Version 2

4.6.3 Digital Twin Framework for Maintenance Version 3

Version 3 of the DT shown below is the most advanced version. The framework has been in use for some time in the manufacturing facility and it has been rolled out across the shop-floor. As shown below, all key assets now have DTs of their key components. This would mean the site has reached a high level of digital maturity.

All key components in key assets are tracked in this version of the DT. There would be potential to expand the DT program to more assets, but a cost benefit analysis would need to be done in-line with the maintenance strategy. The data sets for the key assets is multi-layered in this version of the DT. Primary sensory data is combined with other data sources to allow machine learning algorithms to find patterns to assist the asset. These data sources can include energy data on the asset, OEE, CMM, domain controller data and alarm data. This holistic data pool will increase in value for the asset as more historical data becomes available.

4.6.3.1 Increased data pool

The larger data pool can now be used by the data analytics teams to generate ML models. This work cannot be completed until there is enough data available for the analytics team to train and test models. As previously discussed, the data quality must also be of sufficient standard. The DT now has a prediction tool that can be used for advanced prognostics. The analytics team can use the models that can predict failures to simulate different outcomes for assets based on various conditions.

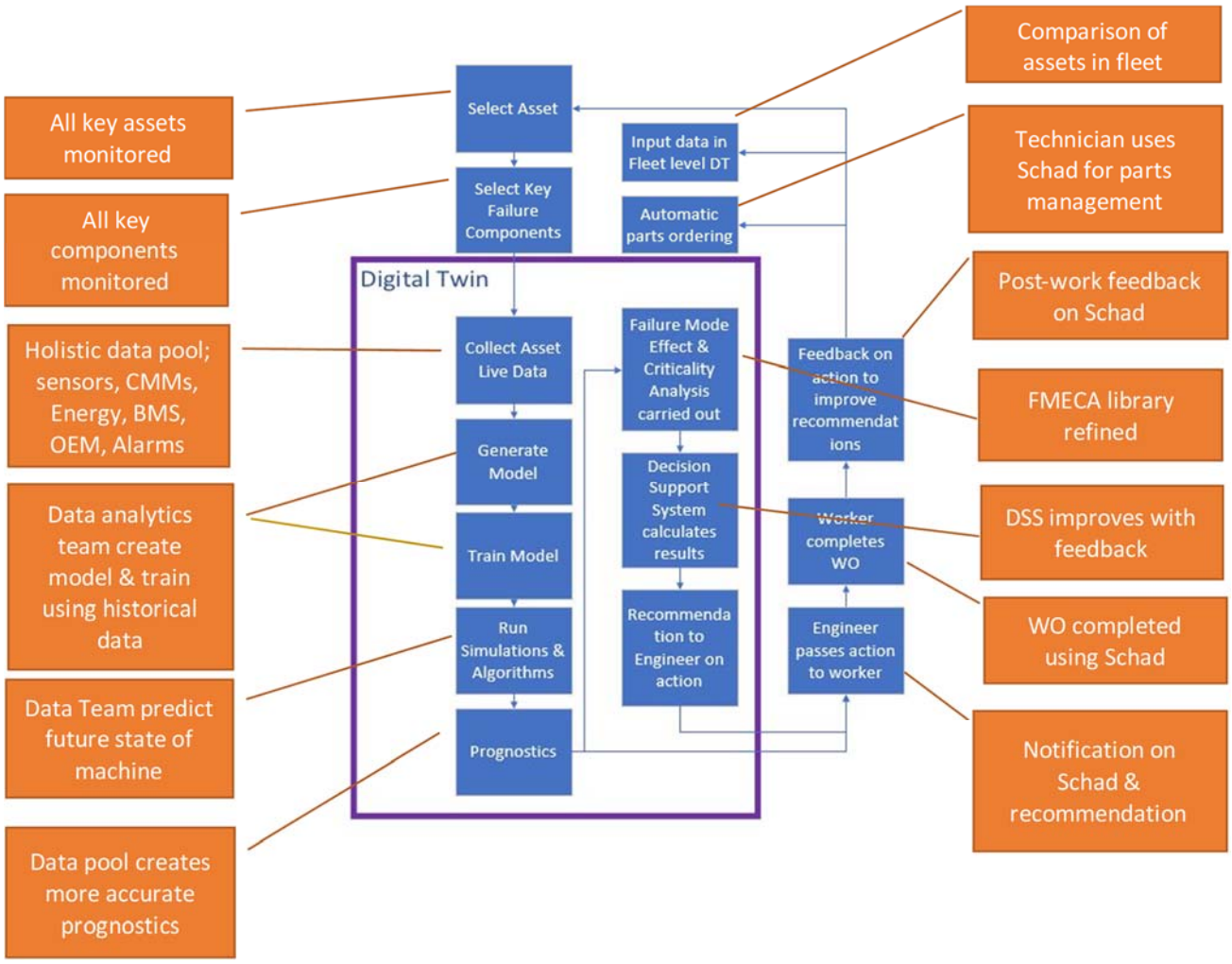


Figure 25 Digital Twin Framework Version 3

The simulation works just discussed has not been completed in the case-study work and it will be sometime before the facility is in a position to do so. There is a misconception in industry management that this work can be completed instantaneously as required. The iterative nature of the DT just discussed shows this ML work is one of the last steps to be implemented, not the first.

The prognostics can be used with the improved FMECA and DSS capabilities to fulfill the CPS aim of having a forward-looking maintenance strategy. The closed feedback loop nature of the framework means the DT can improve as event data is fed back into the DT after work is completed on the asset. This data rich DT can now offer value to similar assets. This is where the idea of a digital fleet comes to the fore. This concept has been discussed in detail previously.

The work so far has looked at the issues that explain why there are so few DT case studies in industry and what are the difficulties in developing the DTs. The framework mentioned above is used to develop a maintenance DT, to understand the true capabilities of the DT, its ideal state must be reviewed to

provide a north star to works towards. The following chapter reviews the process that is fundamental the ideal state DT, and this is machine learning. The chapter reviews the fundamentals of the topic and what parts of the subject could be used to in a DT. This will help to answer the research question of how to build a maintenance DT.

5 Machine Learning & Digital Twin

5.1 Machine Learning Introduction

Machine Learning (ML) is topical at the moment in many industries as it offers the potential to add a competitive edge to a business and leap ahead of competitors. Like any topical subject it can mean different things to different parties and its potential can be overstated in the excitement. At its core machine learning is a method where algorithm search for patterns in data [90].

5.1.1.1 Machine Learning wider use

The above description can be somewhat underwhelming compared with the noise being discussed on social media and online but that in essence is what it can do. People are getting anxious and agitated about it because the potential knowledge and impact ML can have when used in certain areas. The recent Cambridge Analytica case, where data was obtained improperly from Facebook to help presidential campaign teams to build voter profiles is an example of this misguided use [91]. The data was mined and using ML people were classified and targeted to swing votes. This case shows the potential ML has to impact society in a huge way.

It is easy to overstate the potential ML can have in the world if it can be used to decide the next president of a country. What people fail to grasp is that it was being used to complete its primary task; search for patterns in data. Granted in this case it was used in questionable circumstances, people however think ML can be used to do many things away from this core use which is not necessarily true.

ML in other industries such as social media, finance and on the internet can be implemented in a more straightforward manner because of the amount and quality of data available. Although manufacturing facilities produce large data sets, these pale in size when compared with other industries. More data makes ML easier. Manufacturing facilities must be aware that their industry is a unique, complex sector with intricacies that are complex and may not fit seamlessly with topical subjects such as ML.

So, when ML is being introduced into in a large manufacturing facility it is important to temper expectations of staff at all levels as to what it can and cannot do.

5.2 Machine Learning & Industry 4.0

ML is a useful tool for making sense of data and gaining new insights. It is however only the tip of the iceberg in the data work required to see that knowledge. The data preparation and ETL work required beforehand dwarfs the ML test and training steps.

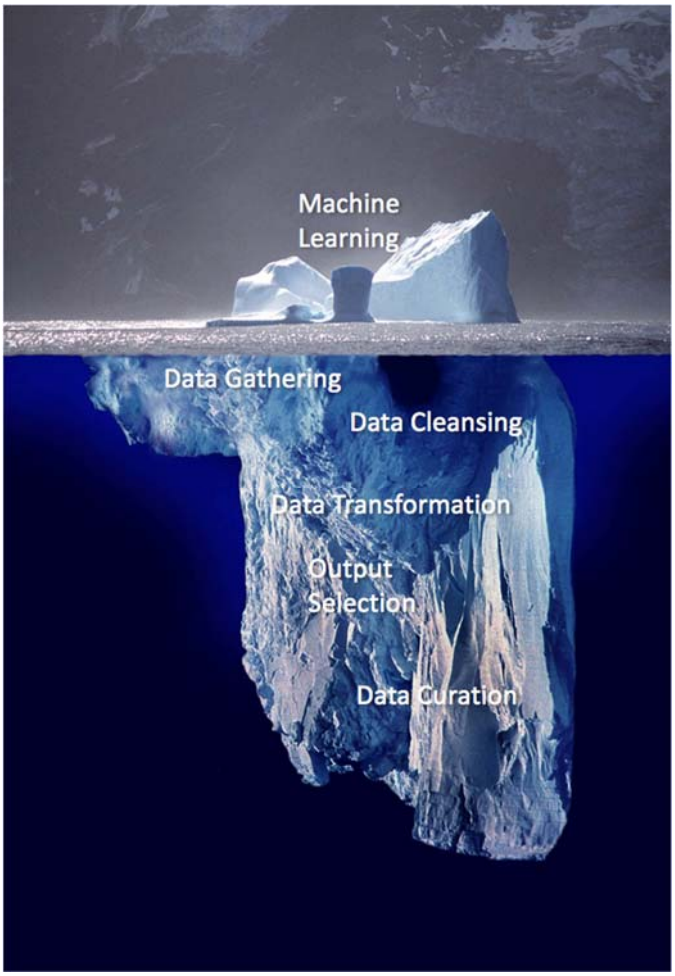


Figure 26 Machine Learning workload is only the tip of the iceberg for data preparation work [92]

The ETL process has been discussed in previous sections, it is important that ML users highlight to the business the time and monetary cost of this work before the benefits of an ML program can be achieved. In the 5C architecture, mentioned before, for CPS in manufacturing facilities ML would lie in levels IV and V. Hence there is a body of work before ML that acts as a foundation upon which the ML system is built.

The manufacturing facility where this case-study has been based have a projects program for digitizing processes. This is split into three sections, Connect, Monitor and Self-Learn. Again, this aligns with the 5C CPS architecture and ML would be in the final bucket of work. It is intuitive that the sensors must be connected and collecting quality data before the ML work can be started.

The desire to implement ML in I4.0 settings is to help drive system performances. ML offers an opportunity to see new insights into the CPS of the facilities, receive this information quicker and reduce human error from processes. Data gathered about assets or systems on the shop-floor is unmined knowledge until it has gone through an ETL process. The complexity of this ETL can vary significantly, from simply displaying data on a dashboard to inputting it into a ML model. ML can be used to generate new knowledge for predictive maintenance, utility optimization and quality control. The push now in industrial analytics is to use historical data gathered with real-time data to make informed proactive decisions quicker. This will help reduce cost, scrap, lost man hours etc.

ML can help business react quicker, moving from diagnostics to prognostics. The human mind is a powerful tool but struggles to analyze large data sets and repetitive tasks well. ML can be used to do these tasks better, freeing people up to solve more complex problems.

5.3 Machine Learning, Algorithms & Functions

Machine learning is a method of using computer algorithm to search for patterns in data. The learning idea discussed here can be broken into three parts; representation, evaluation and optimization. Representation refers to the set of classifiers or identifiers used to categorize the data. Once represented by the different classifiers, each classifier must be ranked for accuracy. This step distinguishes the good and bad classifiers. The final step then is to optimize the search of the classifiers for the highest scoring one. This process is a learning component and hence the name, machine learning.

For machine learning there are different learning concepts, these include supervised learning, unsupervised learning and semi-supervised learning. The work focused on here is supervised learning. With supervised learning the data is labelled so the data has been classified. An algorithm is created based on this classification. Then based of this, prediction of other data is completed using the machine learning algorithms [93].

5.3.1 Machine Learning Sticking Point

Since Machine Learning has started as a concept there have been different sticking points to its successful implementation. These high-level sticking points are memory, data and time. Depending on the time in ML history the current sticking point has changed. In the early use of ML, memory to store all the data would have been a concern and hampered the use of ML. However, technology advancements matching Moore's law rate of improvement have meant in recent times memory is no longer the main concern for ML work. With growing memory capacity, the quantity of data available for the training and test sets became the sticking point. With increased demand and production of hardware and software related to ML the cost of technology to create the data reduced. This reduction meant more data was created and the data sticking point subsided. Now with "unlimited" memory due to large cloud servers and cheap technology to create data sets engineers are swamped with data. The sticking point now is working through all the data that is available and finding the value in it [90].

The sticking point for ML work has changed since its formation as a topic and engineers have solved the issues as they have arisen and will continue to do so. What is recommended when undertaking ML is to use the data to your advantage. Allowing the large quantity of data to inform the correct classification rather than relying on a data scientist to create a smart algorithm is the recommended workflow with ML [90]. The engineers understanding of the data themselves is key to successful use of that data.

5.3.2 Algorithms

ML relies on large data sets to be able to train the models effectively. With large sets of data, it can become messy, not presented in the correct format for analysis and this presents challenges. Also, with big data it can be time consuming to scan the data when completing a search. Hence the make-up of the

algorithm used in the ML process is important as that decides how long the scan takes. This doesn't matter in smaller data sets, but when the datasets contain hundreds of thousandths of rows the scan run time can become too long. The diagram below shows the operation time for "O" time based on different functions. It highlights that the selection of the function type for the ML algorithm is key. The amount of operations required for the algorithm can rise significantly based on the algorithm selected. So it is important to select an algorithm that does not take too long to operate because of the large data sets.

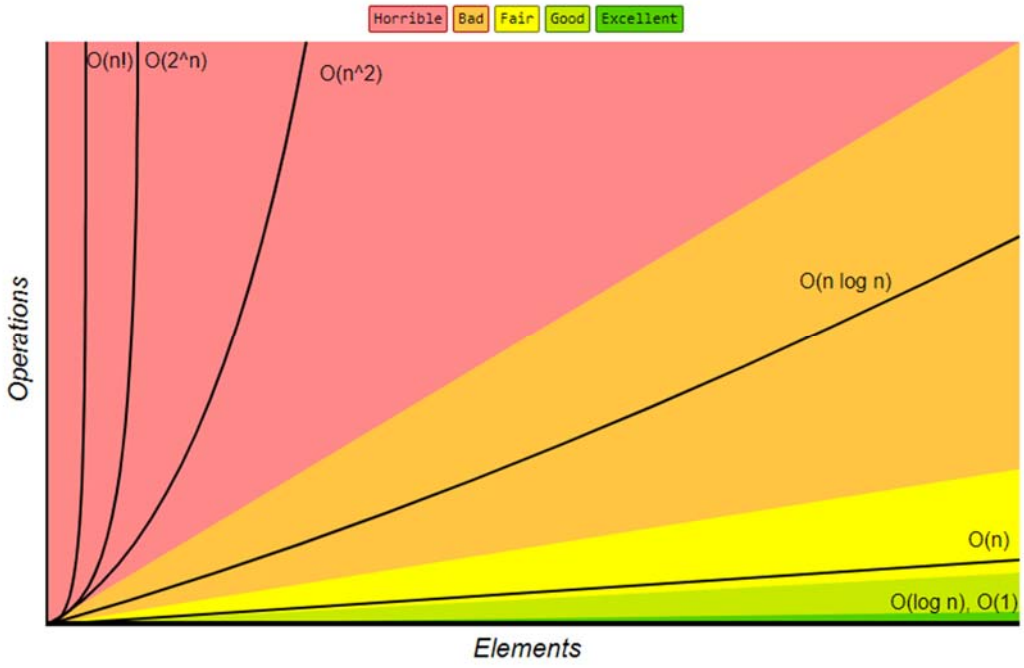


Figure 27 Big O Time Complexity Chart [94]

5.3.3 Data Preparation

For efficient handling of data, we need data structure and design methodologies to ensure the ML algorithms are proficient. To prepare the data various techniques can be used to;

- Sorting
- Graphs
- Greedy algorithms

These methods are not discussed in detail in this document but rather seeks to familiarize the reader with the terms as the ever changing role of a manufacturing engineer moving forward in Industry 4.0 will mean engineers will need to be comfortable discussing ML tasks with data scientists while solving process issues concurrently.

5.3.3.1 Sorting

Divide and conquer techniques like Merge Sort and Randomized Sorting, sort disordered data in a structured manner that can be searched quicker.

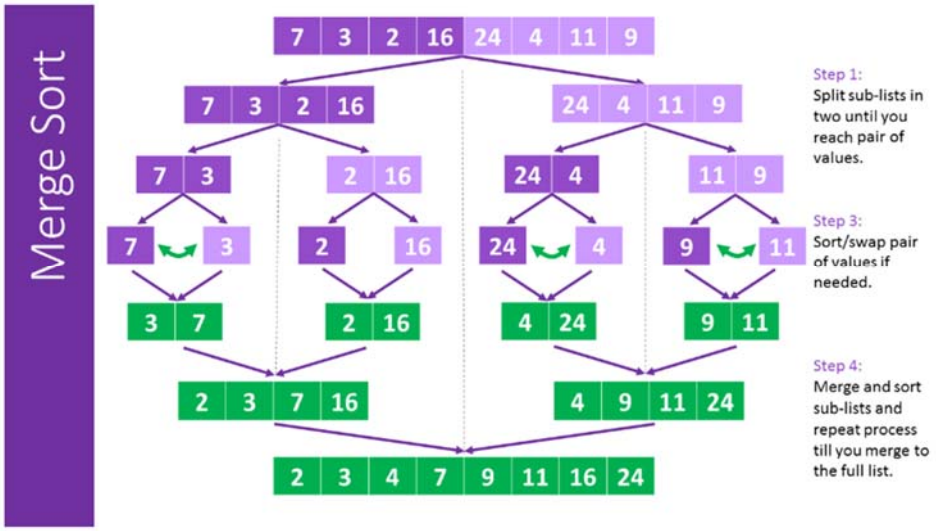


Figure 28 Merge Sort example [95]

Hashing is used to store encrypted data; it creates random but deterministic data. When working with data that may be shared outside the business hashing the data would be important for data security purposes.

5.3.4 Graphs

Graphs are used to model relationships between objects. The relationships can be symmetric or asymmetric. The graphs consist of vertices and edges [90]. Graphs vertices or objects can be any type of information set such as; people, tasks, computers, airports or time. The edges then highlight some connection between the objects; people liking each other, planes that fly to airports or roads between cities.

Like standard tables of data, searching through graphs can be time consuming and must be streamlined to save time. There are two standard algorithms for searching graphs, breadth first search (BFS) and depth first search (DFS). The figure below highlights the different search method implemented by each. BFS is used more for finding the shortest distance, it searches all the neighbors of the node before moving down to the next depth of vertices, such as being used in Google Maps search. Whereas DFS searches from the node as far as possible down one branch before moving onto the next branch, this would be used in scheduling problems.

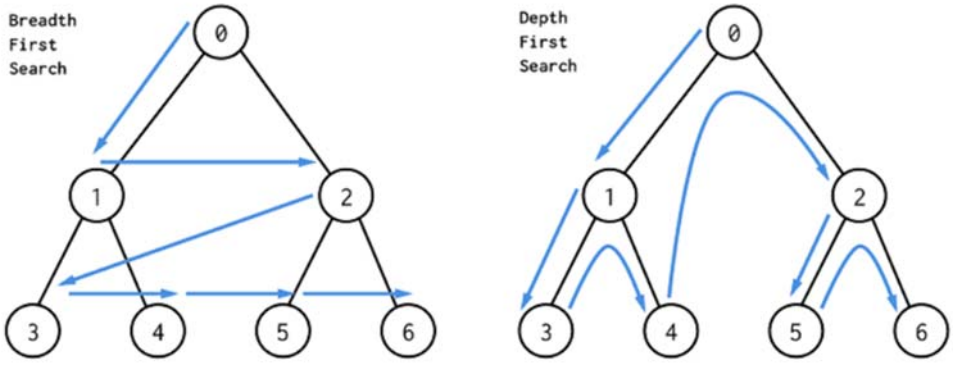


Figure 29 Breadth First Search and Depth First Search [96]

5.3.4.1 Greedy algorithms

Greedy algorithms make locally optimal decisions with the intent of finding the overall optimum decisions. This would be an example of an algorithm technique used in graphs. When the best next immediate vertices are required before moving onto the next layer of vertices this algorithm would be used. The local optimal strategy is selected possibly to the detriment of the best long-term result for the search.

To find the shortest and or fastest possible route on a graph it is necessary to add weights to the edges. This weighted data could be static or live such as traffic congestion on roads. Dijkstra's Algorithm is used to implement this search of graphs, it explores nodes in increasing order of distance [90].

5.4 Supervised Learning Method

Using all the techniques just discussed the data is prepared for use in models. Supervised Machine Learning is the probable ML technique that would be implemented as part of a manufacturing facilities DT process. Although datasets from manufacturing facilities are big, they are not large in big data terms. They can also be quite messy, data from various machine sources will require an amount of ETL work. Supervised learning will ensure the data is prepared correctly and a model of adequate complexity is created to help the engineers predict issues.

The Supervised Learning ML process is shown below. Raw data from the source is ingested, cleaned, patterns are marked, then the data pool is split into training data and test data sets. Once these tests are run the evaluation and optimization of the models is completed.

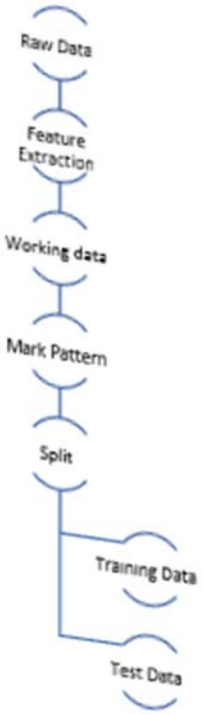


Figure 30 Supervised Machine Learning workflow

5.4.1 Supervise Learning preparation work

Feature extraction involves the collection of measurements and preparation of the data. For a DT this means the infeed of sensor data from the edge to the database location. From here an ETL process will prepare the data for the ML models. Now that the data is prepped it is referred to as working data.

5.4.2 Pattern Marking

Marking the pattern is where the supervision in Supervised Learning is key. The supervisor must determine the features of the input which signify different outputs. This task marks a pattern as a pattern, hence the ML is supervised. It is important that the supervisor lets the computer extract the knowledge of what is a typical pattern it is looking for rather than the human programing the pattern into the algorithm. The computer looks at examples and learns from these examples what a pattern is and what are the typical descriptions of a pattern.

5.4.3 Training and testing

Then the available dataset is split into training and test data sets. With the structure of the ML algorithm selected the training data set should be run on it. Parameters may require selection now for the model to optimize performance. After these adjustments have been made the performance of the ML model is tested on the separate test data set.

5.5 Classification versus Regression

ML learning can be used to calculate discrete and continuous values, these different values in supervised learning are called classification and regression.

- Classification refers to categorizing values with known certain values.
- Regression analysis allows the results to be continuous values.

Classification is used to give a definitive A or B answer such uses include; image classification, identity fraud detection and various types of diagnostics. Regression is based on estimating possible values, predicting and or forecasting, such as population growth prediction, life expectancy and weather forecasting.

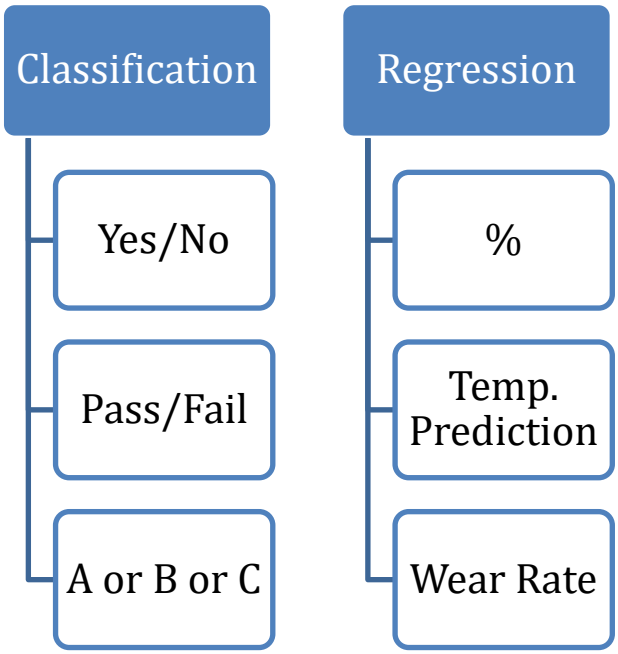


Figure 31 Classification vs Regression

For a DT in a manufacturing facility ML allows captured equipment data to be used to make informed future decisions about the asset. Both classification and regression ML techniques offer potential uses in the maintenance field. Classification can be used to alert workers to when equipment components are in a comprised or failure state.

Regression can be used to predict future states, the models can attempt to predict values for inputs based on data. This future state prediction is a step change in maintenance strategy thinking.

The remainder of this section discusses classifiers that were studied as possible tools that could be installed on the facility to classify failure points.

5.6 Classification

5.6.1 K-Nearest Neighbor

K-Nearest Neighbor is one of the more basic classifiers in ML world, however it offers good potential use in small, less complex data sets from machinery. The algorithm is easy to implement and there is minimal tuning of parameters. The data set is memorized, and previous algorithm techniques are used to find the nearest neighbor to the new point added, as quickly as possible.

5.6.1.1 Selecting K

Looking at the example image below, a new data point is added to the set and based on the Kth adjacent points a decision is made how the value should be classified. A majority vote between the selected points decides the new points classification. Picking the value of K is a bias versus variance task, too few points and it may be inadvertently affected by a stray adjacent point and too many and the selection process is slowed and diluted by all the points far away.

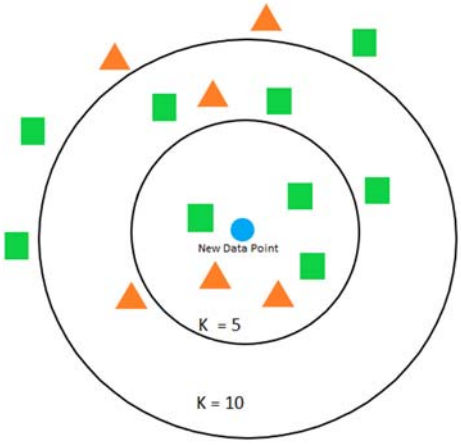


Figure 32 K-Nearest Neighbor Example

Adding dimensions and larger datasets can slow the process and make it a poorer selection for the task. These would not be the foremost problems with machinery datasets.

5.6.1.2 K-Fold Cross Validation

To improve the performance of the algorithm, K-fold cross validation can be used for the data set. This technique splits the training and test data in N sets. Training is completed on one set and the rest are used for testing. Then the process is repeated but the test set is rotated. This is repeated N times, it is a useful procedure where the same overall set of data is reused multiple times to improve the ML accuracy.

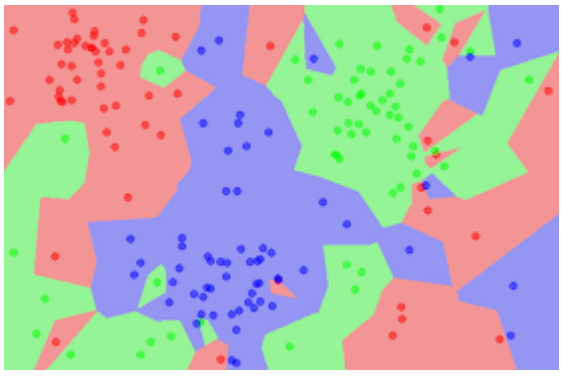


Figure 33 Diagram showing 3 Classes using K-Nearest Neighbor

5.6.1.3 KNN uses

K-Nearest Neighbor can be used to classify points or for regression. For classification it works well with multiple classes, which is not the case with other classifiers such as Support Vector Machine. For regression the Kth values used for in the algorithm forecast the value based on the new inputs. The figure below shows the difference.

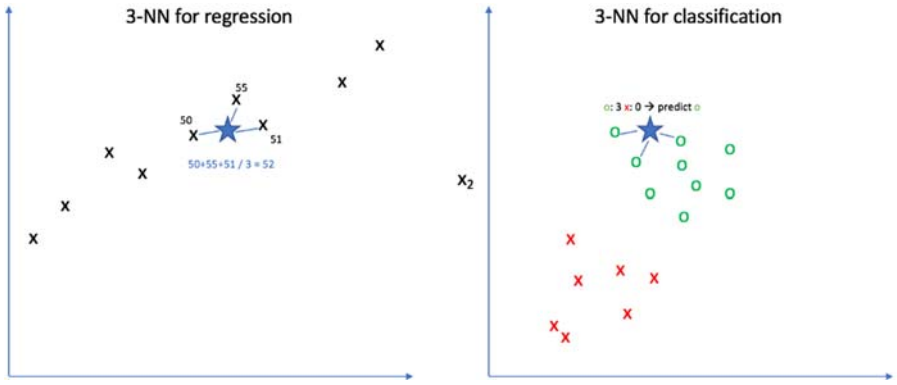


Figure 34 Comparison of K-Nearest Neighbor Regression and Classification

5.6.2 Support Vector Machines

Following on from K-Nearest Neighbor (KNN), Support Vector Machine (SVM) classifier is another popular classifier that could be used in a facilities maintenance DT. Where KNN has the laborious task of considering all points in a data set, SVM creates a boundary around the same classified data points. Then a line can be drawn between the two boundaries and depending on what side of the line a new data point is, decides its classification.

SVM works best when there are 2 classifications and not more. The ideal line that separates the classes maximizes the margin between the classes. This separating line is known as a hyperplane as there may be more than two dimensions involved in the data. The support vectors are then the perpendicular lines to the hyperplane. And depending on a new point's support vector value it can be calculated which class the point is in. SVM has been used by K. Leahy et al. to predict maintenance issues in wind turbines [63].

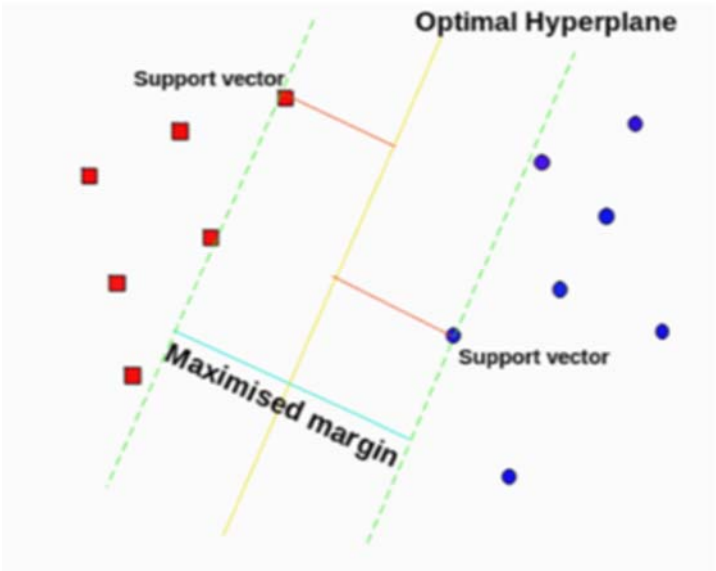


Figure 35 Support Vector Machine Classifier [97]

Data sets in two dimensions can appear visually mixed up and cannot be classified using SVM but adding a third dimension and using a hyperplane separates the classes, as the figure below shows.

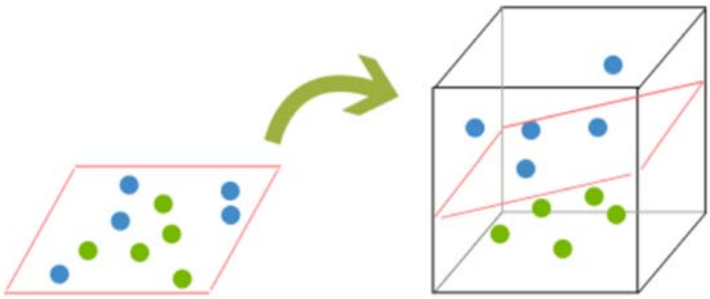


Figure 36 SVM Hyperplane

5.6.2.1 SVM add-ons

For SVM classification there are tools to improve the system, a Kernel trick can curve the SVM line if required to suit the pattern marking better. A cost can be added to the algorithms that impact classification of points near the hyperplane. In some classifications studies a false negative could have large implications, for example for a false negative for medical testing. To prevent this a cost is introduced to the algorithm that reduces this likelihood, as a false positive if highlighted once can be tested again for confirmation whereas a false negative may mean a person lets a disease grow in them unknown to them despite getting tested. So, costing allows overlapping of the hyperplane to improve the classification as required. These tools can be used to improve the SVM, multiple classes can be used by an SVM but it will lead to reduced performance and this must be evaluated by the engineer.

5.6.3 Random Forest Classifier

Another ML method is ensemble learning, this learning method uses multiple learning algorithms, that outperforms those individual algorithms if used on their own. Random Forest and Adaboost are two examples of this learning method, random forest will be discussed in more detail here. The main difference between Adaboost and Random Forest is the difference in how they make the weak learners dependent, Random Forest uses randomization and AdaBoost uses a deterministic strategy [90].

A weak learner is a classifier that has an expected error rate slightly above 50%. A good classifier would have an error rate of 5%. However, combining multiple weak classifiers leads to an accurate result by majority vote. What is important is that the learning algorithms are stochastically independent, this means that the occurrence of one does not affect the probability of occurrence of the other. The Random Forest classifier relies on tree structures, like divide and conquer previously discussed, to classify each point. These multiple trees make up a forest, and the selection of the separation lines on the axes are random and hence the classifier name.

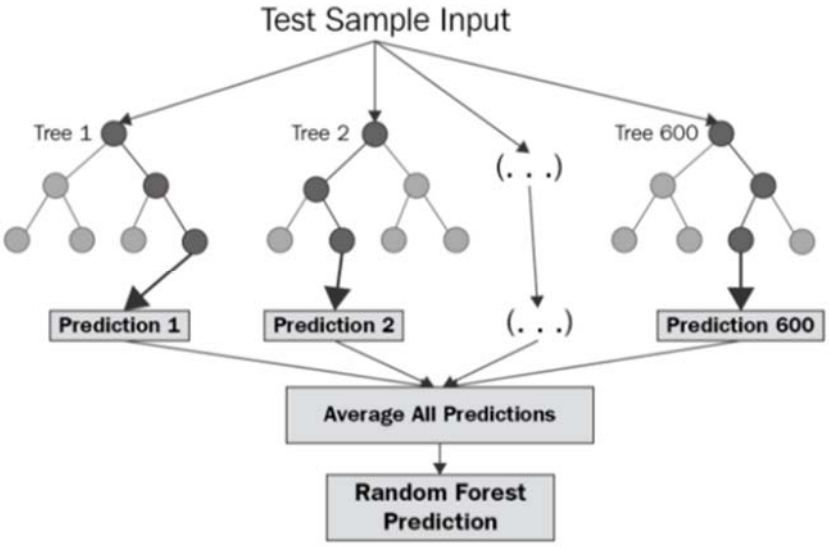


Figure 37 Random Forest Classifier [98]

Above is a figure of Random Forest Classifier. Random Forest is a robust classifier, that doesn't suffer from overfitting and can handle missing data points.

5.6.4 Fitting

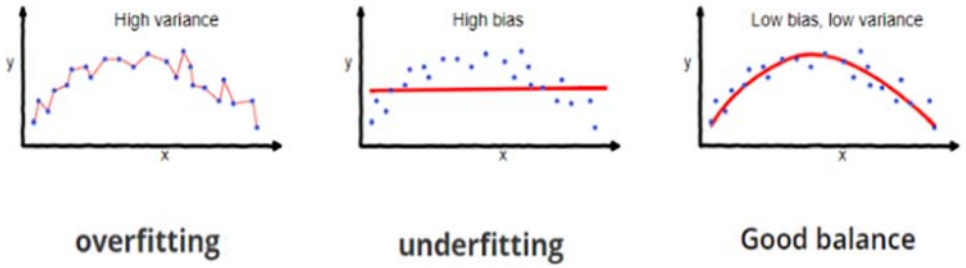


Figure 38 Graphs of ML algorithm fitting with variance and bias error [99]

For supervised ML there is always a balancing act to prediction errors. Overfitting and underfitting can occur because of bias and variance. Overfitting is when the model is too closely marked to the training data. Yes, the model works well for the training data and the training error is low, but the variance is too high once the model is used with the test data. And conversely with underfitting the training error is closer to the test error, but the bias is too high.

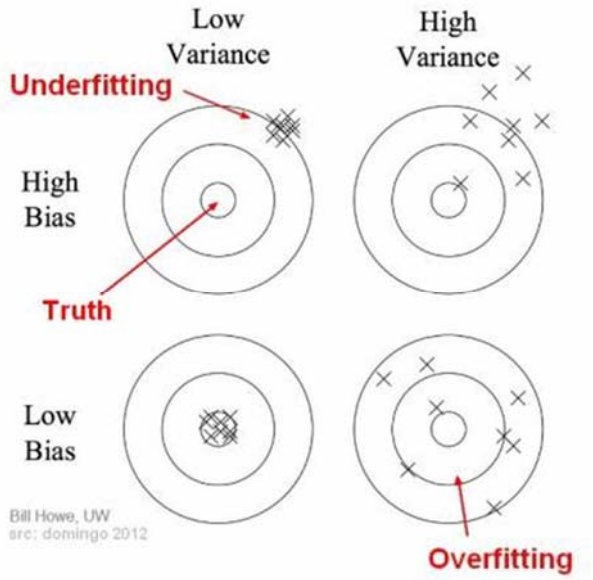


Figure 39 Accuracy mapping due to variance and bias [99]

5.6.5 Machine Learning and Digital Twin

Machine learning is a hot topic across industries at the moment as it is part of an automated business landscape. Its potential to offer value to a manufacturing facility is high. Within that setting it could be used on many data sources such as supply chain, planning and scheduling or for equipment health. To justify its use and cost it must generate information that drives business decisions. This as yet, has not been proven in this case-study and is an area for future work.

For the use of ML in industrial DTs to occur good quality data is required and models that are not prohibitively complicated or expensive must be used. The models discussed above are the most likely candidates to offer this ML solution. Other authors have already used or highlighted Support Vectors as a technique to be used in diagnostics [46, 51, 63]. While this might be the most used technique, each use case must be reviewed on a per case basis to see what format suits the data and desired classification or calculation.

Neural networks have also been mentioned as a ML technique for DTs, this option looks a long way off, especially in regulated industries which are slow to change and alter validated systems [46, 48]. Allowing unsupervised ML techniques to make decisions for equipment is sometime away yet.

Researchers in the work group have started this work to look at ML in industrial settings and there is optimism that the ML work can offer insights that would justify its inclusion in a DT framework like above. The following chapter draws to a close the thesis with discussions and conclusions about the work and the main research questions.

6 Discussion & Conclusions

6.1 Discussion

This body of research looked to examine the use of a DT as part of an improved maintenance strategy. It outlined a framework for developing a maintenance-based DT. Work was done to introduce a DT into the manufacturing facility, it became clear this would not be possible in its entirety currently as the infrastructure, data and analytics capabilities are not in place. This finding is in line with the lack of industrial case-studies in literature [33].

The DT framework created, highlights the projects required to introduce the DT in steps. There are several challenges that must be overcome before the facility sees the benefit of a functioning DT for maintenance. The CBM sensors are the first step in this journey towards a DT. It is through a collection of these projects that the facility will put all the parts of the DT framework in place.

These smaller projects offer less risk than a single large project to introduce a DT immediately. It is intended to commence projects in all areas of the framework; full criticality component review, real-time data analytics, decision support system introduction and failure mode analysis tool.

6.1.1 Significant findings from research

There have been noteworthy findings from the research work;

- It has highlighted the actual difficulties of trying to install a DT in an industrial setting.
- Significant value for a business can be created in a very short timeframe by following steps towards starting a maintenance DT.
- A framework has been created that can be used to guide the future introduction of DTs.

The work has highlighted key challenge's and difficulties that are not discussed in concept papers. The practical work in the facility provided a unique opportunity to test the ability to install a DT. The interviews with key stakeholders highlighted that the work would not be as straightforward as discussed in journal articles. The discussions highlighted a disconnect between academia and the production shop-floor.

The practical work of installing the sensors on the preliminary set of assets highlighted the benefit of digitizing physical assets. The payback on the work was almost instantaneous. This sets a strong precedent for expanding the scope of assets in the program. Business leaders are primarily driven by cost, so it is significant news that the cost benefit of the DT work is so strong.

The difficulties encountered by previous DT work and other I4.0 projects highlighted that a new plan needed to be put in place to increase the probability of success in new DT projects. Reduction in project scale and complexity lead by the DT framework has been the result. The framework creates a clear workflow for workers and links all the necessary parts of a successful maintenance DT together in an effective way.

6.1.2 Comparing results with previous studies

The results from this thesis have highlighted a gap in the research field that only a few researchers have discussed. Academic papers have discussed DTs as a concept, at a distance from the shop-floor [27] [42] [100] [101]. Other articles have installed DTs with much reduced capabilities in the laboratory [40] [36] [14]. All these articles do not show the difficulties of the installation of DT on the production shop-floor. The number of industrial case-studies of DTs is still low which indicates there are still obstacles to its implementation [33].

The gap in research is with regard to case-study examples of DTs being installed in industrial facilities. A body of work in this area will help inform the industry of the success and failure of DTs in production settings.

6.1.3 Limitations of results

The case-study looked at in this thesis is only from one facility, so it is unfair to extrapolate the results from this study for industry in general. However, this facility is part of a larger corporation and no other facility in the corporation is closer to installing a full DT. This fact and the lack of industrial case-study published papers shows the research is at the forefront of the field of study.

The implementation phase of the case-study work was much shorter and not as successful as planned. This constrained the findings possible from the practical work. The practical worktime was curtailed due to administration difficulties in allowing the sensors and analysis platform to be installed in the facility. This can happen in regulated industries which are part of larger corporations that are not as agile as smaller companies. Smaller companies can make changes to facilities quicker and install new technology with less difficulty. With the new technology not fully installed, the Covid-19 pandemic occurred which restricted personnel access to site, this further slowed progress. The disappointing aspect of this is that if the system had been installed on time, it would have worked well in the pandemic as it would have allowed remote monitoring of the critical assets.

6.1.4 Research value with current literature

The research has highlighted the reality of work attempting to create a digital asset. This digital asset needs data to exist, the findings conclude that there is large body of work to create these multiple data pools.

When reviewing vibration data alone, it is a standalone field of engineering. Certain other authors have reviewed DTs in this singular manner and it is apparent that the workload is sizable [36] [40]. The research done here with vibration data reflects this also.

Certain authors have highlighted this, F.Tao et al. comment on some of the challenges facing data-driven smart manufacturing project work such as; lag in data collection technology capabilities, unresolved cloud based analytical issues, such as network unavailability, overfull bandwidth and latency issues [11]. While others have glossed over this workload by commenting that data can be pulled from multiple sources easily. This ETL work needs more case-study examples to accelerate the learnings of practical installation of DTs.

6.1.5 Future research

The future research for this work is focused on the installation of each section of the framework; collect multiple data sources for key assets, train and

run models, full prognostic capabilities, FMECA tool for each asset and a DSS. While this work is being done each part can be iteratively improved.

Other authors have discussed using the DT for wider scopes throughout a business. M. Kunath and H. Winkler discuss using the data for decision support at the system level of manufacturing for order management [102]. The DT can also be used for dynamic scheduling, dynamic calculation of delivery dates and pricing as part of the ordering process and dynamic administration of supply processes [102]. F. Tao et al. also proposes use of the DT in resource management optimization and production planning optimization [101].

F. Tao and M. Zhang investigate the operation and evolution of DTs across a production shop-floor [100]. They propose a shop-floor service system, linking multiple data sources. This is the type of DT that the manufacturing facility tried to generate in 2016. Its ambition is significant, when this can be achieved is unknown as of yet.

L. Hu et al. discuss the use of a cloud-based knowledge resource center for use with DTs [60]. The DTs sit within a cyber-physical cloud manufacturing system. This remote knowledge center mirrors the discussion here about DSS and FMECA libraries.

All these articles discuss similar topics to this thesis, and some are even more ambitious about things that can be achieved by the DT. The work following this case-study will be grounded in practical completion on the production floor where the work can instantaneously add value to the business that own it. And as the DT grows, its value will grow, and it will help bridge the gap between concept DT and realized digital asset.

6.2 Conclusion

This study examined the capacity to install a DT in a large-scale manufacturing facility. The initial review of the topic raised some questions that the case-study work attempted to answer;

- Why are there so few industrial case-studies of digital twins?
- What are the difficulties in constructing a maintenance digital twin?
- How would you build a maintenance digital twin?

When combined the different aspects of work in this thesis have comprehensively answered the three questions above.

The practical work involved collecting tacit knowledge from key stakeholders about the topic in the facility to help answer the questions above and install a maintenance DT as efficiently as possible. The findings from the interview process was that there are a host of issues when trying to complete this I4.0 work in a manufacturing facility. These include business, human, project management and vendors issues.

To negate the impact of as many of these findings as possible a DT framework was created to reduce the project complexity and scale. With the framework in place initial installation work was started. The CBM sensors installed for the DT were an immediate success and they lead business decisions based on their output.

There is still a significant amount of work to be completed for the framework to be filled out. This will take time; the iterative nature of the

framework allows previous parts to work independent of new sections being added.

There is a difficulty in implementing the theoretical idea of a DT to a large-scale manufacturing facility. It is only when one examines previous I4.0 projects do the issues, not covered in concept papers, emerge. These issues provide learnings and once followed the progress towards a DT for maintenance is achievable. Facilities must know that data-based projects in industrial settings pose a considerable challenge and even tougher again in regulated industries. Breaking the CPS goals into smaller projects and obtaining incremental gains means that the I4.0 evolution can succeed.

Future research should consider investigating the use of the framework in other facilities to see the benefit and applicability cross-industry. The payback of the CBM sensors for the DT has been highlighted by this thesis and outlines a model of continued benefit as the work to build a DT is undertaken. Front loading I4.0 project with cost and no payback is not a good working model to be embraced by business.

The contribution of this research has been to bridge the gap between academic studies of DTs for maintenance in industrial settings and real-world applicability. The research has leveraged tacit knowledge from the shop-floor from key stakeholders in this field to apply it to a broader audience using a framework as a guide.

7 References

[1] Medidata, "The Rise of Integrated Data in Medical Devices," Medidata, 2018.

[2] R. Srinivasan and A. Zielinska, "State of the Edge," Seagate, 2019.

[3] D. Gates, E. Gampenrieder, T. Mayor and C. Simpson, "Global Manufacturing Outlook," KPMG International, 2018.

[4] P. Wellener, "Deloitte 2020 Manufacturing Industry Outlook," Deloitte Development LLC, 2019.

[5] J. Manyika, "Manufacturing the future: The next era of global growth and innovation," McKinsey Global Institute, 2012.

[6] A. Zwegers, "International Data Spaces - Artificial Intelligence for Manufacturing," [Online]. Available: www. Internationaldataspaces .org/wp-content/uploads/dlm_uploads/2019/07/20190625-1500-Common-European-Industrial-IoT-by-Arian-Zwegers.pdf . Artificial Intelligence for Manufacturing.

[7] A. Kusiak, "Smart manufacturing," *International Journal of Production Research,* vol. 56, no. 1-2, pp. 508-517, 17 1 2018.

[8] P. O'donovan, K. Leahy, D. O. Cusack, K. Bruton and D. T. J. O'sullivan, "A data pipeline for PHM data-driven analytics in large-scale smart manufacturing facilities".

[9] P. O'Donovan, K. Bruton and D. O'Sullivan, "Case study: The implementation of a data-driven industrial analytics methodology and platform for smart manufacturing," *International Journal of Prognostics and Health Management,* no. Oct 2016, pp. 1-22, 2016.

[10] P. O'Donovan, K. Leahy, K. Bruton and D. T. O'Sullivan, "An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities," *Journal of Big Data,* vol. 2, no. 1, 1 12 2015.

[11] F. Tao, Q. Qi, A. Liu and A. Kusiak, "Data-driven smart manufacturing," *Journal of Manufacturing Systems,* vol. 48, pp. 157-169, 1 7 2018.

[12] S. I. Shafiq, C. Sanin, E. Szczerbicki and C. Toro, "Virtual engineering object/virtual engineering process: A specialized form of cyber physical system for industrie 4.0," in *Procedia Computer Science*, 2015.

[13] S. M. L. Coalition. [Online]. Available: https://smartmanufacturingleadershipcoalition.org/.

[14] W. Yang, K. Yoshida and S. Takakuwa, "Digital Twin-Driven Simulation for a Cyber-Physical System in Industry 4.0 Era," 2017, pp. 227-234.

[15] K. D. Thoben, S. A. Wiesner and T. Wuest, *"Industrie 4.0" and smart manufacturing-a review of research issues and application examples,* vol. 11, Fuji Technology Press, 2017, pp. 4-16.

[16] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee and B. Yin, "Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges," *IEEE Access,* vol. 6, pp. 6505-6519, 13 12 2017.

[17] M. De Villiers, "Three approaches as pillars for interpretive information systems research," *SAICSIT,* no. 2005, pp. 111-120, 2005.

[18] R. Baheti and H. Gill, "Cyber Physical Systems," *The Impact of Control technology,* no. 2011, pp. 161-162, 2011.

[19] K. M. Alam and A. El Saddik, "C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems," *IEEE Access,* vol. 5, pp. 2050-2062, 2017.

[20] I. Graessler and A. Poehler, "Intelligent control of an assembly station by integration of a digital twin for employees into the decentralized control system," in *Procedia Manufacturing*, 2018.

[21] R. Stark, C. Fresemann and K. Lindow, "Development and operation of Digital Twins for technical systems," *CIRP Annals - Manufacturing Technology,* no. 1918, p. 4, 2019.

[22] V. Rudtsch, J. Gausemeier, J. Gesing, T. Mittag and S. Peter, "Pattern-based business model development for cyber-physical production systems," in *Procedia CIRP*, 2014.

[23] Y. Cai, B. Starly, P. Cohen and Y. S. Lee, "Sensor Data and Information Fusion to Construct Digital-twins Virtual Machine Tools for Cyber-physical Manufacturing," *Procedia Manufacturing,* vol. 10, pp. 1031-1042, 2017.

[24] L. Monostori, B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn and K. Ueda, "Cyber-physical systems in manufacturing," *CIRP Annals,* vol. 65, no. 2, pp. 621-641, 2016.

[25] J. Lee, B. Bagheri and H. A. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems," *Manufacturing Letters,* vol. 3, pp. 18-23, 1 1 2015.

[26] E. Glaessgen and D. Stargel, "The Digital Twin Paradigm for Future NASA and US Air Force Vehicles," *American Institute of Aeronautics and Astronautics,* no. 53rd Structures, Structural Dynamics, and Materials Conference.

[27] M. Grieves, "Digital Twin: Manufacturing Excellence through Virtual Factory Replication," Michael Grieves, 2015.

[28] B. Schleich, N. Anwer, L. Mathieu and S. Wartzack, "Shaping the digital twin for design and production engineering," *CIRP Annals - Manufacturing Technology,* vol. 66, no. 1, pp. 141-144, 2017.

[29] G. N. Schroeder, C. Steinmetz, C. E. Pereira and D. B. Espindola, "Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange," *IFAC-PapersOnLine,* vol. 49, no. 30, pp. 12-17, 2016.

[30] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, Springer International Publishing, 2016, pp. 85-113.

[31] M. Macchi, I. Roda, E. Negri and L. Fumagalli, "Exploring the role of Digital Twin for Asset Lifecycle Management," *IFAC-PapersOnLine,* vol. 51, no. 11, pp. 790-795, 1 1 2018.

[32] E. Negri, L. Fumagalli and M. Macchi, "A Review of the Roles of Digital Twin in CPS-based Production Systems," *Procedia Manufacturing,* vol. 11, pp. 939-948, 2017.

[33] W. Kritzinger, M. Karner, G. Traar, J. Henjes and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine,* vol. 51, no. 11, pp. 1016-1022, 1 1 2018.

[34] i-Scoop, "i-Scoop," [Online]. Available: https://www.i-scoop.eu/internet-of-things-guide/industrial-internet-things-iiot-saving-costs-innovation/digital-twins/.

[35] S. Boschert, C. Heinrich and R. Rosen, "Next Generational Digital Twin," *TMCE,* no. 2018, pp. 209-218, 2018.

[36] S. Haag and R. Anderl, "Digital twin – Proof of concept," *Manufacturing Letters,* vol. 15, pp. 64-66, 1 1 2018.

[37] Q. Qi, F. Tao, Y. Zuo and D. Zhao, "Digital Twin Service towards Smart Manufacturing," in *Procedia CIRP*, 2018.

[38] R. Söderberg, K. Wärmefjord, J. S. Carlson and L. Lindkvist, "Toward a Digital Twin for real-time geometry assurance in individualized production," *CIRP Annals - Manufacturing Technology,* vol. 66, no. 1, pp. 137-140, 2017.

[39] R. Rosen, G. Von Wichert, G. Lo and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," in *IFAC-PapersOnLine*, 2015.

[40] M. Vathoopan, M. Johny, A. Zoitl and A. Knoll, "Modular Fault Ascription and Corrective Maintenance Using a Digital Twin," *IFAC-PapersOnLine,* vol. 51, no. 11, pp. 1041-1046, 1 1 2018.

[41] Bottani E, Cammardella A, Murino T and Vespoli S, "From the Cyber-Physical System to the Digital Twin: the process development for behaviour modelling of a Cyber Guided Vehicle in M2M logic".

[42] M. Ayani, M. Ganebäck and A. H. Ng, "Digital Twin: Applying emulation for machine reconditioning," in *Procedia CIRP*, 2018.

[43] D. Issacs, A. A. Astarola and J. A. B. Diaz, "Making Factories Smarter through Machine Learning," *IIC Journal of Innovation,* no. 3, pp. 29-40, 2017.

[44] V. Wowk, "Machinery Vibration: Measurement and Analysis," New York, McGraw-Hill Education, 1991, p. 8.

[45] J. Wang, Y. Ma, L. Zhang, R. Gao and D. Wu, "Deep learning for smart manufacturing; Methods and applications," *Journal of Manufacturing Systems,* no. 48, pp. 144-156, 2018.

[46] A. K. Jardine, D. Lin and D. Banjevic, *A review on machinery diagnostics and prognostics implementing condition-based maintenance,* vol. 20, 2006, pp. 1483-1510.

[47] B. C. Menezes, J. D. Kelly, A. G. Leal and G. C. Le Roux, "Predictive, prescriptive and detective analytics for smart manufacturing in the information age," in *IFAC-PapersOnLine*, 2019.

[48] S. Zhang and R. Ganesan, "Multivariable Trend Analysis using Neural Networks for intellignet diagnositcs of rotating machinery," *ASME,* vol. 119, pp. 378-384, 1997.

[49] U. D. o. Energy, "Operations & Maintenance Best Practice: A Guide to Achieving Operational Efficiency release 3.0," Pacific Northwest National Laboratory, 2010.

[50] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori and M. Helu, "Cloud-enabled prognosis for manufacturing," *CIRP Annals - Manufacturing Technology,* vol. 64, no. 2, pp. 749-772, 2015.

[51] Y. Peng, M. Dong and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: A review," *International Journal of Advanced Manufacturing Technology,* vol. 50, no. 1-4, pp. 297-313, 9 2010.

[52] F. Tao, H. Zhang and A. N. A. Liu, "Digital Twin in Industry; State-of-the-Art," *IEEE Transactions on Industrial Informatics,* vol. 15, no. 4, pp. 2405-2415, 2018.

[53] F. Longo, L. Nicoletti and A. Padovano, "Ubiquuitous knowledge empowers the Smart Factory; The impacts of a service oriented Digital Twin on entreprises performance," *Annual Reviews in Control,* no. 2019, 2019.

[54] F. Tao, Q. Qi, L. Wang and A. Y. Nee, "Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison," *Engineering,* vol. 5, no. 4, pp. 653-661, 1 8 2019.

[55] J. Morgan and G. E. O'Donnell, "Multi-sensor process analysis and performance characterisation in CNC turning—a cyber physical system approach," *International Journal of Advanced Manufacturing Technology,* vol. 92, no. 1-4, pp. 855-868, 1 9 2017.

[56] J. Downey, D. O'Sullivan, M. Nejmen, S. Bombinski, P. O'Leary, R. Raghavendra and K. Jemielniak, "Real Time Monitoring of the CNC Process in a Production Environment- the Data Collection & Analysis Phase," in *Procedia CIRP*, 2016.

[57] Y. Xu and M. Ge, "Hidden Markov model-based process monitoring system," *Journal of Intelligent Manufacturing,* no. 16, pp. 337-350, 2004.

[58] T. Mukherjee and T. DebRoy, "A digital twin for rapid qualification of 3D printed metallic components," *Applied Materials Today,* vol. 14, pp. 59-65, 1 3 2019.

[59] D. Botkina, M. Hedlind, B. Olsson, J. Henser and T. Lundholm, "Digital Twin of a Cutting Tool," in *Procedia CIRP*, 2018.

[60] L. Hu, N. T. Nguyen, W. Tao, M. C. Leu, X. F. Liu, M. R. Shahriar and S. M. Al Sunny, "Modeling of Cloud-Based Digital Twins for Smart Manufacturing with MT Connect," in *Procedia Manufacturing*, 2018.

[61] G. L. Knapp, T. Mukherjee, J. S. Zuback, H. L. Wei, T. A. Palmer, A. De and T. DebRoy, "Building blocks for a digital twin of additive manufacturing," *Acta Materialia,* vol. 135, pp. 390-399, 15 8 2017.

[62] P. Stavropoulos, A. Papacharalampopoulos, E. Vasiliadis and G. Chryssolouris, "Tool wear predictability estimation in milling based on multi-sensorial data," *International Journal of Advanced Manufacturing Technology,* vol. 82, no. 1-4, pp. 509-521, 1 1 2016.

[63] K. Leahy, C. Gallagher, K. Bruton, P. O'Donovan and D. T. O'Dullivan, "Automatically Identifying and Predicting Unplanned Wind Turbine Stoppages Using SCADA and Alarms System Data: Case Study and Results," in *Journal of Physics: Conference Series*, 2017.

[64] H. Han, B. Gu, T. Wang and Z. R. Li, "Important sensors for chiller fault detection and diagnosis (FDD) from the perspective of feature selection and machine learning," *International Journal of Refrigeration,* vol. 34, no. 2, pp. 586-599, 3 2011.

[65] A. K. Garga, K. T. Mcclintic, R. L. Campbell, C.-C. Yang, M. S. Lebold, T. A. Hay and C. S. Byington, "Hybrid Reasoning for Prognostic Learning in CBM Systems".

[66] M. Jackson, Understanding Expert Systems: Using Crystal, Wiley 1st Edition, 1986.

[67] T. H. Uhlemann, C. Schock, C. Lehmann, S. Freiberger and R. Steinhilper, "The Digital Twin: Demonstrating the Potential of Real Time Data Acquisition in Production Systems," *Procedia Manufacturing,* vol. 9, pp. 113-120, 1 1 2017.

[68] J. L. O. Soldatos and F. Cavadini, The Digital Shopfloor; Industrial Automation in the Industry 4.0 Era, Gistrup: River Publishers, 2019.

[69] "Webopedia," [Online]. Available: https://www.webopedia.com/TERM/E/ETL.html.

[70] A. Reeves, "Managing data in Motion," Morgan Kaufmann, 2013, p. 29.

[71] "Oracle," [Online]. Available: https://docs.oracle.com/cd/B19306_01/server.102/b14223/ettover.htm. [Accessed 29 07 2020].

[72] E. Negri, L. Fumaalli, C. Cimino and M. Macchi, "FMU-supported simulation for CPS Digital Twin," *Procedia Manufacturing,* no. 28, pp. 201-206, 2019.

[73] A. Felsberger, B. Oberegger and G. Reiner, "A review of decision support systems for manufacturing," *iKnow,* p. 2016, 2016.

[74] T. Vafeiadis, D. Kalatzis, A. Nizamis and D. Ioannidis, "Data analysis and visualization framework in the manufacturing decisoin support system of COMPOSITION project," *Procedia Manufacturing ,* no. 28, pp. 57-62, 2019.

[75] A. Jardine, "Optimzing Condition Based Maintenace Decisions," *Proceedings Annual Reliability and Maintainability Symposium,* pp. 90-97, 2002.

[76] M. Pirog-Mazur, "The model of decision support system for a manufacturing company," *Artificial Intelligence Driven Solutions to Business and Engineering Problems,* pp. 46-53.

[77] D. Bumblauskas, D. Gemmill, A. Igou and J. Anzengruber, "Smart Maintenance Decision Support Systems based on coporate big data analytics," *Expert Systems with Applications,* no. 90, pp. 303-317, 2017.

[78] P. Girdhar, "Pratical Machinery Vibration Analysis and Predictive Maintenace," Oxford, Elsevier, 2004.

[79] "Cyber Physics," [Online]. Available: https://www.cyberphysics.co.uk/topics/waves/superposition.htm.

[80] "Azima DLI," [Online]. Available: http://azimadli.com/vibman/summaryofamplitudeunits.htm.

[81] A. Fernandez. [Online]. Available: https://power-mi.com/content/study-vibration.

[82] Siemens. [Online]. Available: https://community.sw.siemens.com/s/article/root-mean-square-rms-and-overall-level.

[83] Erbessd, "Erbessd," [Online]. Available: https://www.erbessd-instruments.com/docs/phantom/phantom-sensors/vibration-phantom.

[84] "Origin Lab," [Online]. Available: https://www.originlab.com/doc/Tutorials/2D-Waterfall.

[85] A. Services. [Online]. Available: https://www.ariesmar.com/pms/vibration-analysis.html.

[86] B. Lynch, "Lindskog Balancing," [Online]. Available: https://www.lindskog.com/2014/08/motor-shaft-alignment/.

[87] "Vibsens," [Online]. Available: http://www.vibsens.com/iso10816-charts/.

[88] "Cincinnati test Systems," [Online]. Available: https://www.cincinnati-test.com/pressure-decay-test/ultrasonic#:~:text=Ultrasonic%20leak%20detection%20utilizes%20high,if%20the%20leak%20is%20turbulent..

[89] "Flir," [Online]. Available: https://www.flir.com/discover/professional-tools/how-to-detect-a-water-leak-with-thermal-imaging/.

[90] C. University, "Edx," Edx, [Online]. Available: https://www.edx.org/course/machine-learning-for-data-science-and-analytics.

[91] N. Confessore, "The New York Times," [Online]. Available: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

[92] A. Freed, "Freedville," [Online]. Available: https://freedville.com/blog/2017/03/05/machine-learning-is-just-the-tip-of-the-iceberg-5-dangers-lurking-below-the-surface/.

[93] S. Direct, "Science Direct," [Online]. Available: https://www.sciencedirect.com/topics/computer-science/supervised-learning.

[94] Bigocheatsheet, "Bigocheatsheet," [Online]. Available: https://www.bigocheatsheet.com/.

[95] 101computing, "101computing," [Online]. Available: https://www.101computing.net/merge-sort-algorithm/.

[96] D. Zaltsmna, "Dev.to," [Online]. Available: https://dev.to/danimal92/difference-between-depth-first-search-and-breadth-first-search-6om.

[97] Mitosis, "Mitosis," [Online]. Available: https://www.mitosistech.com/support-vector-machine/.

[98] C. Bakshi, "Random Forest Regression," 2020. [Online]. Available: https://levelup.gitconnected.com/random-forest-regression-209c0f354c84.

[99] S. Singh, "Towards Data Science," [Online]. Available: https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229.

[100] F. Tao and M. Zhang, "Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing," *IEEE Access,* vol. 5, pp. 20418-20427, 24 9 2017.

[101] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *International Journal of Advanced Manufacturing Technology,* vol. 94, no. 9-12, pp. 3563-3576, 1 2 2018.

[102] M. Kunath and H. Winkler, "Integrating the digital twin of the mamnufacturing system into a decision support system for improving the order management process," *Procedia CIRP,* no. 72, pp. 225-231, 2018.

[103] J. Kunthong and M. Konghirun, "IOT based traction motor drive condition monitoring in electric vehicles part1," *IEEE PEDS,* 2017.

[104] B. A. Talkhestani, N. Jazdi, W. Schloegl and M. Weyrich, "Consistency check to synchronize the Digital Twin of manufacturing automation based on anchor points," in *Procedia CIRP*, 2018.

[105] C. Li, S. MahaDeVan, Y. Ling, S. Choze and L. Wang, "Dynamic Bayesian network for aircraft wing health monitoring digital twin," *AIAA Journal,* vol. 55, no. 3, pp. 930-941, 2017.

[106] A. A. Malik and A. Bilberg, "Digital twins of human robot collaboration in a production setting," in *Procedia Manufacturing*, 2018.

[107] J. Guo, N. Zhao, L. Sun and S. Zhang, "Modular based flexible digital twin for factory design," *Journal of Ambient Intelligence and Humanized Computing,* vol. 10, no. 3, pp. 1189-1200, 13 3 2019.

[108] T. D. West and M. Blackburn, "Is Digital Thread/Digital Twin Affordable? A Systemic Assessment of the Cost of DoD's Latest Manhattan Project," in *Procedia Computer Science*, 2017.

[109] B. Scaglioni and G. Ferretti, "Towards digital twins through object-oriented modelling: a machine tool case study," *IFAC-PapersOnLine,* vol. 51, no. 2, pp. 613-618, 1 1 2018.

[110] İ. Erozan, "A fuzzy decision support system for managing maintenance activities of critical components in manufacturing systems," *Journal of Manufacturing Systems,* vol. 52, pp. 110-120, 1 7 2019.

[111] L. Wang, M. Törngren and M. Onori, "Current status and advancement of cyber-physical systems in manufacturing," *Journal of Manufacturing Systems,* vol. 37, pp. 517-527, 1 10 2015.

Appendix A: Journal Articles reviewed per sub-heading

Table 15 Papers reviewed

Reference	Level of Integration	Setting	Topic	Maintenance Applicability	Data Pool
R. Stark [21]	DT	Laboratory	Assembly Cell	Medium	Wide
J. Kunthong [103]	DS	Laboratory	Motor	Medium	MC Specific
J. Morgan [55]	DS	Laboratory / Production	CNC	Medium	MC Specific
J. Downey [56]	DS	Laboratory / Production	CNC	Medium	MC Specific
J. Lee [25]	DT	Theory	CNCs	Medium	Wide
R. Gao [50]	DT	Theory	All MCs	High	Wide
B.A. Talkhestani [104]	DT	Theory	All MCs	Low	Wide
C. Li [105]	DM	Theory	Aircraft wing	High	Statistical detail
Y. Peng [10]	DS	Theory	All MCs	High	Wide
A. Jardine [46]	DS	Theory	All MCs	High	Wide
F. Tao [54]	DT	Theory	All MCs	Low	Wide
Y. Xu [57]	DS	Laboratory / Production	CNC	High	MC Specific
M. Kunath [102]	DT	Theory	All MCs	Low	Wide
P. O Donovan [8]	DS	Theory	All MCs	Medium	Wide
E Negri [72]	DS	Laboratory	Mobile phone assembly	Low	Handful of O/Ps
F.Tao [52]	DT	Theory	All MCs	Medium	Wide
F.Tao [101]	DT	Theory	All MCs	Medium	Wide
T. Mukherjee [58]	DM	Theory	3D Printing	Low	MC Specific
M. Macchi [31]	DT	Theory	All MCs	Low	Wide
Q. Qi [37]	DT	Theory	All MCs	Low	Wide
B Schleich [28]	DM	Theory	All MCs	Low	Wide
D Botkina [59]	DS	Laboratory	Cutting Tool	Medium	MC Specific
F. Tao [100]	DT	Theory	All MCs	Medium	Wide
H. Han [64]	DS	Laboratory / Production	Chiller	High	MC Specific
M Vathoopan [40]	DT	Laboratory / Production	Mechatronic component	High	MC Specific
L. Hu [60]	DS	Laboratory / Production Focused on the cloud	3D Printing	Low	MC Specific
A.A. Malik [106]	DS	Laboratory	Cobot	Low	MC Specific
S. Haag [36]	DT	Laboratory	Bending Bench	Low	MC Specific
W. Yang [14]	DT	Laboratory	Miniature Vehicle	Low	MC Specific
I. Graessler [20]	DM	Laboratory	Assembly Station	Low	MC Specific
G.L. Knapp [61]	DS	Laboratory / Production	3D Printing	Low	MC Specific
G.N. Schroeder [29]	DS	Laboratory / Production	Valve	High	MC Specific
J. Guo [107]	DM	Theory	Factory Design	Low	Wide
K. Leahy [63]	DS	Laboratory / Production	Wind turbine	High	MC Specific
M. Grieves [30]	DT	Theory	All MCs	Low	Wide
T.D. West [108]	DT	Theory	All MCs	Low	Wide

6.2 Conclusion

T.H.J. Uhlemann [67]	DT	Theory	All MCs	Low	Wide
R. Soderberg [38]	DS	Theory	Assembly Line	Low	Wide
M Ayani [42]	DM	Laboratory	Assembly Line	Low	MC Specific
Y. Cai [23]	DS	Laboratory / Production	Cutting Tool	Medium	MC Specific
R.Rosen [39]	DT	Theory	Assembly Line	Low	MC Specific
B Scaglioni [109]	DS	Laboratory / Production	Cutting Tool	Medium	MC Specific
S. Boschert [35]	DS	Laboratory	Train line sensors	Medium	MC Specific
P. Stavropoulos [62]	DS	Laboratory / Production	Cutting Tool	Medium	MC Specific

DT - Digital Twin
DS - Digital Shadow
DM - Digital Model