A design-based study of data-driven asset management for Storm Surge Barriers

Implementing predictive maintenance, digital twins and realizing data governance for the current asset management at Rijkswaterstaat

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Management of Technology**

Faculty of Technology, Policy and Management

by

Ir. Hamza Azad

Student number: 4505349

To be defended in public on April 8th 2022

Graduation committee

Chairperson	: Prof.dr.ir. Marijn Janssen, Section ICT
First Supervisor	: Prof.dr.ir. Marijn Janssen, Section ICT
Second Supervisor	: Dr.ir. Bert Enserink, Section PA
External Supervisor	: Yigal Levin MSc, Rijkswaterstaat
External Supervisor	: Rens van Maarschalkerweerd MSc, Rijkswaterstaat

Preface

First of all, I would like to thank my supervisors, Marijn Janssen, Bert Enserink, Yigal Levin, Rens Maarschalkerweerd and Bernhard Thieme, for supervising me through this challenging but also exciting thesis.

Moreover, many accompanied me through my school carrier. Firstly, my HBO-bachelor group: Haci Fatih, Mohammed, Kubilay, Floda and Botan. We really had a great time during classes and made the food places near our school and mosques temporarily wealthy.
Secondly, my friends made at TUDelft. I really enjoyed the struggling time together, particularly during the bridging program when the stakes were high.
The last group are my previous employers, who gave me the financial means to pay for my education and provide for my family. Marcel van Heusden employed me at Blokker and gave me opportunities to grow my social skills and work ethic. Jorn van Putten, after completing my internship, gave me a position at OCI. My first step as an engineer. Sowande Boksteen was my last employer, who financed my cost for this master.

My family: my sisters and my mother. I would like to thank them for their support and care, especially after our father passed away more than a decade ago.
Finally, I would thank and praise Allah (God) for this experience once in a lifetime. Without you, I would not be here where I am today.

Hamza Azad

Rotterdam, Feb 17, 2022

Executive summary

Rijkswaterstaat's Storm Surge Barriers are expensive assets providing economic value with the expectation of future returns. The maintenance is executed by the asset management department, which aims at maximizing operability without sacrificing safety and reliability. However, regular maintenance becomes more stringent as circumstances change, and a budget shortfall is expected in the same timeframe. These operational and economical interferences urge the asset management department to adapt using more data-driven approaches replacing (some) physical inspections and only conducting maintenance when needed. This so-called data-driven asset management has the potential to reduce costs and ensure reliability.
Despite the promise, there is no guidance for developing data-driven asset management. In this thesis, design principles are developed to assist in better decision-making using data-driven technologies, resulting in effective maintenance. These new design principles integrate three different areas into the current asset management: Digital Twins (DT), providing a virtual environment for safely testing in various scenarios, Predictive Maintenance (PdM) for confidently predicting a future asset failure and Data Governance (DG), ensuring the data quality for appropriate decision-making for the mentioned data-driven technologies. In other words, the goal of this research is to develop design principles that overcome real-world asset data-related problems during the implementation of DT, PdM and DG within the asset management context.

The qualitative design-based research methodology consists of the following steps: Identifying real-world problems, developing appropriate solutions, and then finalizing by prescribing actionable sentences, also known as design principles. This research design was demonstrated in a real-life case for the Storm Surge Barriers asset class at Rijkswaterstaat (RWS). Two data-collecting research methods were utilized. The first data collecting method was the literature study, which acquired information from 25+ scientific papers from three developing fields in the scientific literature: PdM, DT and DG. The second was a series of semi-structured interviews held with 33 RWS interviewees in different parts of the organization, collecting real-life asset data problems and the associated best practices. To guide the interviews across the data-driven developments and the current status of data governance within RWS, an interview protocol was constructed. The interviews results revealed contrasting views on (1) the SSB openness to new technology, (2) the alignment between asset management priorities and the data-driven technology, and (3) data access by third parties. These contrasting viewpoints unknowingly construct invisible (almost) impermeable walls between different layers in the RWS organization, which prevents knowledge spillovers, resulting in departments maturing at different rates and impeding the understanding and communication between them.

As part of the research design, four types of design principles were formulated. The first was the interview-derived design principles. The second type was the literature derived design principles acquired by the results from the literature study. After that, the interview and the literature-derived principles were combined into 'the refined design principles’. The last type is the data governance design principles, for which the problems were inspired by semi-structured interviews and are solely solved by the current data governance literature.

Thereafter, the refined and data governance design principles were tested within the context of the DT and the PdM by using data flow diagrams. The benefit of using the data flow diagrams is to test if the selected design principles improve the practices of asset management.

Consequently, 33 design principles were developed across the four types of design principles that prescribe guidance on how PdM, DT and DG contribute to developing data-driven asset management. From these, 4 were data governance design principles, and 12 were refined design principles. The refined design principles were not found in prior literature and could be further grouped into short-term and long-term relevance for asset management.
The novelty that these refined design principles bring is twofold, filling in two gaps in the literature. Firstly, It brings clear guidelines which were previously scattered and unclearly presented across the scientific literature. Secondly, these design principles are approached in a novel manner by combining the core insights from the literature and the inclusion of empirical best practices in a real-life case.

The results of implementing the relevant data governance and refined design principles into the AM are twofold. Firstly, the DeP’s enhanced the DT from primarily visualization into a more multifunctional usable digital asset for various asset management purposes. Secondly, implementing the same set of principles enhanced the PdM by including the most important factors to produce reliable automatic predictions to better manage the maintenance schedule.

The managerial recommendation is that the digital twins and predictive maintenance development need to develop hand in hand with data governance aspects and utilize a system to centralize the knowledge across all storm surge barriers, like periodic communities of practices. A lack of these two crucial developments will inhibit RWS long-term vision of transitioning from a pure civil organization to a hybrid organization: A synergy between civil engineering discipline and data-driven technologies.

List of abbreviations

AM	Asset Management
CBM	Condition based maintenance
CoP	Communities of Practice
CPPS	Cyber-physical production system
DB	Database
DBR	Design-Based Research
DDP	Data-driven Projects
DeP	Design Principles
DFD	Data Flow Diagram
DG	Data Governance
DG-DeP	Data governance design principle
DT	Digital Twin
ERP	Enterprise Resource Management
Int-DeP	Interview-derived design principle
IQ	Interview Questions
Lit-DeP	Literature-derived design principle
MoC	Management of Change
MoT	Management of Technology
OS	Operating System
OTS	Operator Training Simulator
PdM	Predictive maintenance
PLC	Programmable Logics Controllers
Ref-DeP	Refined design principle
RQ	Research Questions
RUL	Remaining Useful Lifetime
RWS	Rijkswaterstaat
SSB	Storm Surge Barrier
Sub-RQ	Sub research questions

Table of Contents

List of figures

List of tables

1. Introduction

This chapter introduces the thesis by describing that operational and economical interferences result in asset management to adapt more data-driven approaches. From this, the research objectives are derived. Subsequently, the research questions are defined based on the research objectives. Thereafter, the research's case selection and relevance are discussed in three parts: The scientific, managerial and the MSc Management of Technology (MoT) relevance. Finally, the thesis outline briefly describes the organization of the coming chapters to answer the research questions successfully.

1.1. Research problem

Assets are expensive resources that provide economic value which are owned, controlled and maintained by a person, company or state with the expectation of future returns (Adam Barone, 2022). Every type of organization using these assets to support their core tasks has one common problem: maximizing operability without sacrificing safety and reliability. This is owing to the fact that as an asset is used or sits idle for an extended period of time, it degrades, which is an unavoidable occurrence. For these, regular inspections of the assets are needed. This challenge can be further mitigated using systematic approaches and best practices mostly executed by the Asset Management (AM) department. A well-integrated AM can affect each part of an organization, growing uptimes, lowering costs, and improving firm reputation. If reliable information and powerful decision support tools are integrated within the AM, then the maintenance, repair, and renewal costs are dramatically reduced. A well-integrated AM is generally achieved by decades of experience where the external factors are readily constant (Faiz & Edirisinghe, 2009).

Regular inspections of assets become more stringent as circumstances change, i.e., increasing sea levels due to climate change, unexpected pandemic, and stricter regulations urging the AM to adapt. In other words, these novel changes that were previously seen as an 'constant variable' need to be offset by adapting the current AM. One key area that can facilitate the AM to adapt to is to use 'data as an asset' which supports reducing costs and ensuring reliability (Chen et al., 2021). Data as an asset refers to data-driven approaches in which sensors are used to collect data about the quality of the assets. In this way replacing physical inspection and ensuring that maintenance is only conducted when needed.

Therefore, the 'data as an asset' core tasks for AM should be to provide information at the right time, in the right format, in front of the right people, at the right request, and at the right level (Faiz & Edirisinghe, 2009). This can be addressed by effectively incorporating technologies that manage asset data and assist the decision-makers with data-driven asset insights.

Two prevalent data-driven technologies relevant for the AM are digital twins (DT) and Predictive Maintenance (PdM). The DT provides a virtual environment in which the asset can be tested in various scenarios and the PdM confidently predicts a future asset failure. These two technologies can assist in combination with the asset manager's knowledge and experience to perform better decision-making, resulting in data-driven asset management.

In addition, an essential factor using 'data as an asset' for AM decision-making is the 'quality of data'. The data quality affects the effectiveness of any IT-related work. The quality of the data used to create reports and make decisions can only be as good as the data itself. Although data-driven AM has many promises, the issues surrounding data quality or lack of quality blocks this. This even is worsened because (1) data is spread and scattered in various systems within an organization, (2) data is collected, sustained, and used by various organizational levels, and (3) most organizations did not include 'the quality of data' within their data management systems. Appropriate data governance (DG) can address the above-mentioned data quality issues. It is possible to ensure adequate data quality by establishing a DG program. The DG program gives data managers the authority to monitor data quality, especially in the case of utilizing PdM and DT as these collect (big) data, maintain it, and assist decision-making(Kuan Cheong et al., 2007).

The current status in academia and industry is that there is no guidance for developing data-driven asset management by incorporating DT, PdM and DG. The literary background will further elaborate upon the lack of no guidance within sections 2.2.3 and 2.2.4.
As a result, in this thesis, new design principles (DeP) are developed and integrated three different areas into the current asset management: DT, providing a virtual environment for safely testing in various scenarios, PdM for confidently predicting a future asset failure and DG, ensuring the data quality for appropriate decision-making for the beforementioned data-driven technologies, resulting in effective maintenance.

1.2. Research objective

As a consequence of no guidance for developing data-driven AM, the following general and specific objectives are formulated for incorporating new broad fields such as PdM, DT and DG in the current AM.

General objective: *To identify design principles that overcome the encountered problems during implementation of DT, PdM and DG within the asset management context.*

These design principles can provide broader applicable directions useful for academia and industry and improve the adoption of data-driven AM.

Specific objectives:

1. To assess the encountered problems, and best practices from various ongoing Data-driven Projects (DDP) related to implementing DT, PdM and data governance for the AM purposes.
2. To prioritize the found best practices by including the SSB demand for AM purposes.
3. Combining the Interview-derived DeP's (Int-DeP) and literature-derived DeP's (Lit-DeP) to form a generic and wider applicable Refined DeP (Ref-DeP) and data governance DeP's (DG-DeP).
4. To implement the ref-DeP's and DG-DeP's into a Data Flow Diagram (DFD) to find opportunities and bottlenecks between the interactions of the applicable DeP's.

1.3. Research questions

Following the general and specific research objectives, the following research questions have been developed accordingly:

RQ1: What are the design principles for asset management?

1.1. What are the encountered problems hampering the DT and PdM implementation?

1.2. What does SSB asset management demand from DT, PdM and data governance?

1.3. What are the best practices for the encountered problems and SSB demand to implement for DT and PdM?

1.4. What are the current data governance practices for DT and PdM implementation?

The core purpose of the RQ1 is *to develop* a set of standalone design principles that prescribed how PdM, DT and DG *contribute* to the current AM. In continuation with these formulated design principles, RQ2 will *relate* only the applicable standalone design principles within a DFD-analysis to show how it is implementable for the current AM.

RQ2: How can the design principles be used for asset management?

2.1. What do the DFD-DT and DFD-PdM look like, based on the design principles?

2.2. How can data governance design principles be implemented within the DFD-DT and DFD-PdM to make the usability of PdM and DT more effective for AM purposes?

2.3. What are the opportunities and bottlenecks found from the constructed DFD?

1.4. Case selection

This paper addresses the research based on an actual case study at the governmental organization Rijkswaterstaat (RWS). RWS manages most, if not all, civil assets in the Netherlands. These are broken down into three asset classes: bridges (1000+), tunnels (20+) and 6 Storm Surge Barriers (SSB) which facilities to efficient infrastructure and protection from high sea levels and stormy weather. With this amount of assets, RWS is the largest asset manager in the Netherlands.
The fit between the research and RWS case is as follows:

The asset management activities will increase in the coming decades because of the increased infrastructure usage. Simultaneous in the same time frame, a budget shortfall is expected. Therefore, to offset it, technologies like DT and PdM will be used as leverage to perform AM in a more uniform and model-based way instead of in the past employing ad-hoc and documented-based working methods. The client confined this research to investigate the implementation of data-driven technologies for the AM of the SSB asset class.

1.5. Relevance of this research

This section highlights this research's scientific, managerial, and MoT-related relevance.

1.5.1. Scientific relevance

This study adds to the growing body of knowledge in the field of data governance. Because this topic is still in its infancy, there are numerous significant holes to be filled. This led to the most present data governance research being generic and abstract (Schranz et al., 2020).
In specific, one of the major holes is the lack of data governance principles (guidelines) for the barely researched AM sub-field. This is due to the fact that AM has distinct characteristics and therefore the current literature deriving data governance principles do not explicitly contribute.
As a result, the academic contribution of this research is twofold. Firstly, a set of data governance design principles tailored for AM are formulated. These are derived by combining the RWS empirical case study and the current data governance literature. Secondly, from the DFD-analysis the insights (bottlenecks and opportunities) are extracted by incorporating the data governance principles to two novel AM-related technologies: DT and PdM. These two contributions (data governance design principles and insights) will contribute to further developing the current status of this burgeoning field of interest.

1.5.2. Managerial relevance

The long-term vision is for RWS to transition the SSB asset class from a pure civil organization to a hybrid organization by implementing information technology within the current civil asset management. This is accomplished in three steps, for which this study contributes primarily to the first two-layer.
Firstly, gather information on whether everyone at RWS understands each other, the motivations, and what must be done in the coming years. As a result, this research will provide DeP derived from interviews discussing DT and PdM initiatives within RWS.
Secondly, develop a vision and make choices regarding the usage of information technology at the SSB in the short and long term. Here, the DFD will encompass the refined DeP providing RWS with a systematic way to have a better grasp on strategizing and make decisions to approach the socio-technological problem.
Lastly, the implementation phase consists of an iterative execution plan, starting with data visualization/analysis, adopting the DT's, and covering the organizational aspects needed to implement change. In this report, the implementation is not covered. However, insight will be prescribed in the recommendation section.

1.5.3. MoT relevance

To be eligible to graduate from the MSc Management of Technology (MoT) program, this thesis should meet three general criteria to indicate a typical 'MoT master thesis'.

Firstly, the research should be involved with a high-tech situation. Here, the situation is to develop from the current ad-hoc and documented-based AM to a more data-driven AM facilitating data-driven decision support. Therefore, cutting edge technologies are utilized and data governance is realized. With this technological incorporation for an outdated way of AM the first criteria are met.

Secondly, it should explain the technology as a corporate resource. In this case, the PdM and DT are technologies used to improve the current SSB asset management. Therefore, an understanding is

created by provided DeP's that are prescriptive sentences providing a better grip on implementing these novel technologies within the asset management context.
Lastly, the report should use scientific techniques to analyze problems. Here, qualitative data collecting, and data analysis methods are used from MOT2312 (Uma Sekaran, 2018). A prevalent data analysis within this thesis is the DFD-analysis which is found outside the compulsory MoT curriculum. The purpose of the DFD-analysis is to relate the current asset management, two new technology and data governance. By using multiple appropriate scientific methods it satisfies the last criteria to be relevant a typical 'MoT master thesis'.

1.6. Thesis outline

The report is structured as follows:

- **Chapter 2**: This chapter provides the literature background used as the basis for the rest of the thesis. First, the major abstract concepts defined in various ways are discussed, and definitions are presented, which will be used for this report. Next, a brief explanation is given of how the literature prescribes to use Design-Based Research (DBR) and links it with design principles. Finally, several PdM, DT and GM case studies are analyzed to identify problems and the associated case solutions.
- **Chapter 3**: The research methodology elaborates on the research design, which consists of a qualitative approach. Then, the methods used for data collection and analysis are expanded upon to develop design principles and how these design principles facilitate creating data-driven asset management.
- **Chapter 4**: In this chapter, the results of the semi-structured interviews are given. The interviews are transcribed, followed by an analysis of the Data-Driven Projects (DDP) to formulate Int-DeP's and describe the current data governance status at RWS.
- **Chapter 5**: Then, the Lit-DeP's are formulated and combined with the Int-DeP's to formulate a set of Ref-DeP's focusing on PdM and DT implementation. Moreover, literary solutions are provided for the current DG problems faced by RWS to form GM-DeP's.
- **Chapter 6**: Here, the DFD-PdM-DG and DFD-DT -DG are built and integrated with the relevant ref-DeP's and DG-DeP's to identify how the DeP's can be used for the asset management at RWS.
- **Chapter 7**: Finally, the report's conclusions are summarized, remarks on the findings are made, and the study's limitations are mentioned in this chapter. Furthermore, a reflection on the MoT program is provided, and recommendations are made for future research.

2. Literature background

This chapter provides the literary background to the rest of the thesis. First, the major abstract concepts prone to be interpretable are discussed to one final definition that will be used for this report. Hence, the final definitions can be used for the empirical data collecting, and analysis. Next, a brief literary explanation is given about the research methodology and its relationship with design principles. The purpose is to reliable develop the research methodology which will result in a strategy for producing design principles. Finally, several PdM, DT and DG literature case studies are analyzed to identify problems and the associated case solutions, which are in some way applicable within the AM context. These literature case studies identify novelties from empirical design principles or enhance it with well-known processes from literature.

2.1. Literature study

For the empirical data collection (i.e., interviews, surveys) new trending technologies and broadly defined working methods need to be discussed. However, these concepts are abstract and have multiple interpretations. For this reason, these abstract concepts are defined using current scientific literature to keep the respondents on the topic leading to increased validity of the data collected. The four concepts that will be covered are asset management, predictive maintenance, and data governance.

2.1.1. Asset Management (AM)

A general description of asset management is cited by Brous et al.: 'Asset management is generally understood to be the set of activities of a business objective associated with: identifying what assets are needed; identifying funding requirements; acquiring assets; providing logistic and maintenance support systems for assets; and disposing or renewing assets to effectively and efficiently meet the desired objective' (Brous et al., 2019).This is a broad definition and applicable to various industries. Nevertheless, AM should be understood and contextualized within the case study: SSB, which are governmental assets protecting from flooding.
Hence, Bart Vonk et al. quotes AM within the SSB context as: 'A systematic approach to ensure that throughout the whole life cycle of an asset, or group of assets, the investment costs, the performance achieved, and the risks faced are balanced' (Vonk et al., 2020). This definition is sharper described as it mentions the 'risks factor'. In the 'SSB world', risks are known as the 'fail chance estimation', one of the most important parameters. For this sole reason, the definition proposed by Vonk et al. is used in this thesis.

2.1.2. Predictive Maintenance (PdM)

Asset management is managed by using various maintenance strategies, which is the main topic of this section. All the maintenance methods are introduced, and the position of predictive maintenance is justified within the maintenance tree.

In literature, many maintenance strategies are found. In this literature study, the categorization and characterization by Errandonea et al. are followed as he generated general definitions by analyzing over 167 papers and due to their relevancy by publishing in 2020 (Itxaro Errandoneaa, 2020).

There are five known maintenance strategies to handle maintenance activities across various industries:

Reactive maintenance: Other names are corrective maintenance or failure-based maintenance. These unscheduled emergency repairs are event-driven. Some examples are unexpected breakdowns or leakages which were not observed beforehand. This maintenance strategy is generally the most expensive, up to 3x more costly than preventive maintenance. The highest cost related to reactive maintenance is high spare part inventory, low production availability, high machine downtime, high overtime labour costs & dependency on vendor equipment for immediate delivery of spare parts. This maintenance strategy is only used for non-critical assets with lower impact failures (Mobley, 2002) (Itxaro Errandoneaa, 2020).

Preventive maintenance: The synonyms of this strategy are scheduled maintenance, time-driven maintenance, or time-based maintenance. The reactive maintenance strategy can be significantly reduced by utilizing preventive maintenance. This method relies on the original equipment manufacturers' estimated equipment lifetime and the experiences of the asset managers to plan maintenance activities over periods preventing asset breakdowns and reducing plant downtime.
This is the most used strategy as it is straightforward, result-driven and convenient to plan annual maintenance costs. However, it is not optimal for non-time driven failures accounting for 89% of all failures, as shown in figure 2.1. Also, 'over maintenance' of assets that may be operated 2 or 3 years leads to more unnecessary maintenance costs. The following maintenance strategies can significantly reduce non-time driven failures and 'over maintenance'.

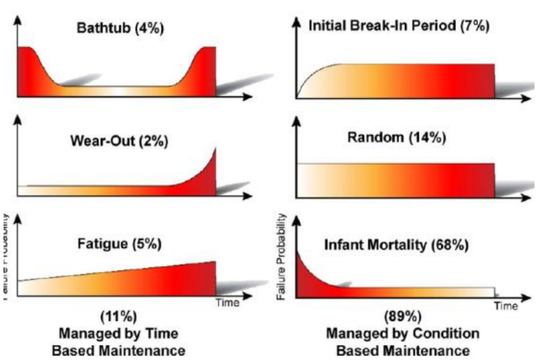


Figure 2. 1: left-hand side is the time-based failures accounting for 11% of the causes. The right-hand side moves past the time-based failure accounting for 89% of the total failures (Hashemian, 2011).

Condition-based maintenance (CBM): A different term is diagnostic-based maintenance. This strategy monitors asset behaviours, deviations and degradation based on its current Condition using IoT to plan maintenance activities accordingly. For instance, the pump's vibration, in and out pressure and flow rate can be continuously measured to observe changes for the same parameters. This strategy is suitable for non-time driven failures as this account for most of the failures, as seen in figure 2.1.
These diagnostics can be advanced using artificial intelligence to find implicit data in the following maintenance strategy.

Predictive maintenance: Another name for this strategy is prognosis maintenance. This maintenance strategy uses all data in the system and its surroundings, called big data and merges using mathematics (statistics and probability techniques) numerical or physical models to predict the remaining asset lifetime. This approach is called a data-driven or model-driven approach. This way, the degradation of the asset is diagnosed with parameters unknown for CBM (i.e. corrosion, velocity profiles, and so on forth) with high reliability to plan better data-driven maintenance activities.

Prescriptive maintenance: Also known as knowledge-based maintenance. At this last stage, the maintenance based on prediction is being optimized. Here, the model prescribes the maintenance activities with high reliability resulting in the optimal service, cost and safety to the same extent possible by using all the available technologies.

The assessment for which of the five maintenance strategies are used is found in figure 2.2. The first subdivision is if the maintenance strategy is planned. If it is not, a reactive maintenance strategy is followed. Otherwise, it is one of the remaining four, which are all proactive maintenance strategies. Afterwards, a series of characteristics need to be identified to categorize which maintenance strategy is executed, as shown in figure 2.2.

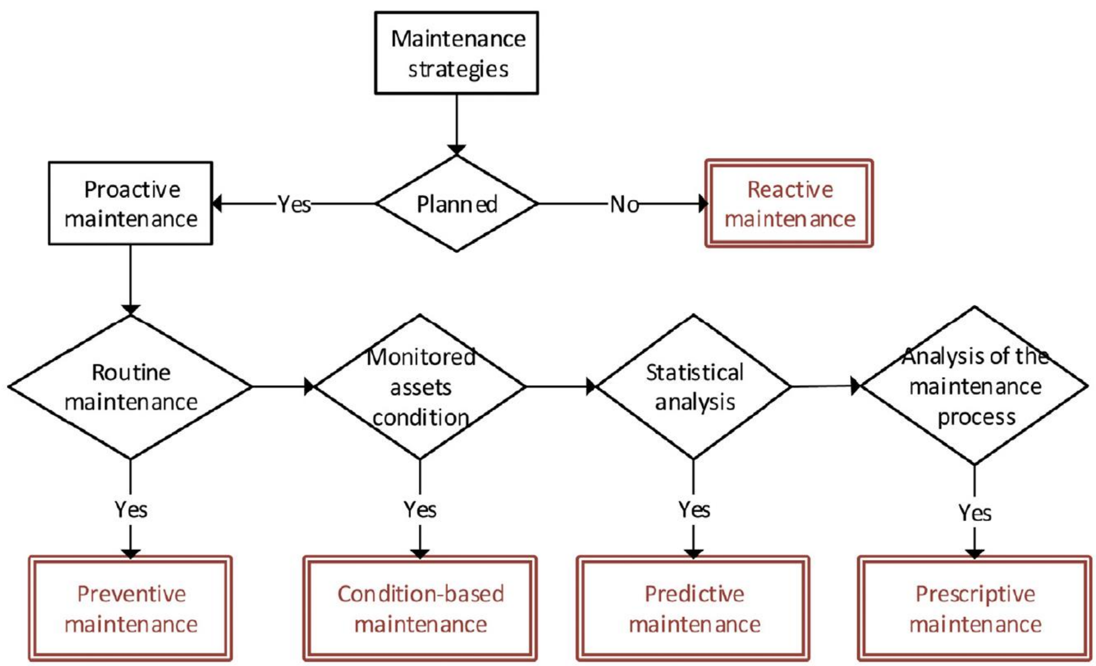


Figure 2. 2: The decision tree categorises all known maintenance strategies (Itxaro Errandoneaa, 2020).

2.1.3. Digital Twins (DT)

Asset management is next to (predictive) maintenance strategies associated with DT, covering this section. Firstly, the term 'DT' is defined. Secondly, the phases of DT are categorized. Finally, the relationship between DT stages and all the maintenance strategies are elaborated.

From literature, a variety of DT definitions found started from 2012. There were even some that were hugely incorrect (Itxaro Errandoneaa, 2020) (Zheng Liu, 2018). Thus, a general definition was found by researching 169 papers related to DT and postulated the following:

A digital twin consists of three parts: the real physical asset or system, the digital/virtual replica, which is an integrated multi-physics, multi-scale, probabilistic model, and the relationship between them being fully automated data flows, sending and receiving (Itxaro Errandoneaa, 2020).

All three parts, the physical asset, virtual model of the physical asset, and the communication between them, predict the environmental responses and unexpected events on the assets and systems, like leakages. This helps the control and decision-making before the event occurs, especially to maintenance activity to the extent of the remaining asset lifetime.

2.1.4. Data governance (DG)

In this section, the concept of data governance is covered extensively. The term data governance is defined. Subsequently, the use of a framework is elaborated with five interrelated data domain decisions. Afterwards, this leads to three approaches to execute governance on the data. Finally, the concept of data stewardship, part of data governance, is introduced.

Organizations, especially in the public sector, make an impactful decision based on collected, combined and used large amounts of various data sets fragmented over departments boundaries. However, the lack of control over the data flows creates uncertainties and makes the data-driven decision less trustworthy and riskier. To ensure a high degree of fit-of-use (meta) data, organizations turn to data governance to exert control over the data quality (M. Janssen, 2020).

Data governance is defined in the literature in various ways. three of them are mentioned below:

- 'The process by which a company manages the quantity, consistency, usability, security and availability of data' (Chang, 2007).
- 'Organizations and their personnel defining, applying and monitoring the patterns of rules and authorities for directing the proper functioning of, and ensuring the accountability for, the entire lifecycle of data and algorithms within and across organizations' (M. Janssen, 2020).
- 'Data governance refers to who holds the decision rights and is held accountable for an organization's decision-making about its data assets' (Brown, 2010).

At the core of every definition is that the importance of data governance is to define policies and procedures ensuring proactive and effective data management (Chang, 2007).

2.2. Design principles (DeP)

Design principles are a natural and logical outcome of design-based research. Therefore, the relationship between Design-based research and DeP's are briefly defined and explained in 2.2.1 and 2.2.2. Consequently, from the PdM and DT literature, problem-solutions are enumerated in section 2.2.3. and the data governance problem-solution in section 2.2.4.

2.2.1. Defining Design-Based Research (DBR)

Design-based research (DBR) is defined as a methodology focusing on improving learning by developing the appropriate solutions to real-world problems (Herrington et al., 2009).
In general, DBR requires more than demonstrating a particular design or refuting others designs. It also

needs the researcher to move beyond a specific design and generate evidence-based claims about learning that combines current theoretical issues and empirically derived knowledge.

The benefit of utilizing DBR is its importance in sharing and distributing findings and concepts (Wang & Hannafin, 2005). These are valuable because of the contribution they offer to the professional community on how to get a better grip on facing real-world problems.
This methodology fits especially to socio-technological, which deals with approaching a complex organizational design that interacts with humans and technology (Herrington et al., 2009).

The outcomes from the DBR are DeP's. These connect encountered problems, instructions, theories of learning, which advances practical and theoretical understanding (Herrington et al., 2009).

2.2.2. Defining DeP

Design principles express in a fashion that can be used to prescribe a widely applicable practice or heuristic (Wang & Hannafin, 2005). Design principles are best articulated in active language and therefore applied quickly when faced with similar situations or problems (Wang & Hannafin, 2005).

To simplify it for experts and designers, DeP's are frequently provided in the form of a list of criteria for certain learning environments and outcomes and frequently begin with a verb. Additionally, DeP's are not fixed like rules, but advice on how they might benefit from the results of past empirical experiences and current scientific literature (Herrington et al., 2009).

2.2.3. Fields within DT and PdM

From the PdM and DT literature there is no explicit guidance for developing data-driven asset management. However, by critically analyzing the case studies and identifying problems and the associated case solutions it provides some potential applications within the AM context. Hence, 9 problem-resulting case studies were identified. These will serve in further chapters by identifying novelties from the empirical design principles or enhance them with the 9 problems-resulting case studies.

2.2.3.1. Problem-resulting case study 1: *Knowledge creation in multidisciplinary teams*

Problem

For post-industrial societies, knowledge is one of the most important resources for an organization. This is especially for a dynamic environment where new problems arise that need to be faced using 'continuous creating knowledge'. Such is achieved by bringing various engineering disciplines together for a synergic solution to specific problems. However, managing multiple teams across the firm introduces redundant projects, concentrated knowledge within one division and minimal knowledge transparency (Fong, 2003).

Resulting case study

Fong proposed a model identifying five steps of knowledge development and their interdependent relationships to manage multiple teams, which is depicted in figure 2.3.
The problem was formulated as a 'knowledge sharing' problem, but it encompassed four other knowledge processes that are entwined and found across every project. Therefore, to incorporate the

boundary-crossing, knowledge generation, knowledge integration and collective project learning when managing knowledge within an origination, division or even within a program (Fong, 2003).

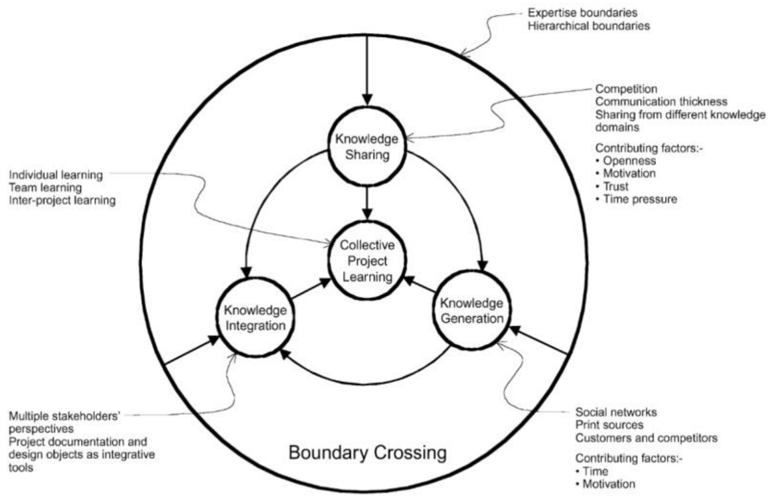


Figure 2. 3: the interdependent relationships between 5 components to create and manage knowledge within a firm. (Fong, 2003)

2.2.3.2. Problem- resulting case study *2: Integrate an adequate information exchange between the shopfloor and enterprise Resource Planning (ERP) system*

Problem

Data-driven technologies like DT have the potential to revolutionize AM. However, the existing solutions are disconnected from enterprise systems and restricted to individual applications. The worst-case scenario could result in time-consuming manual synchronization.
This could be improved by combining and integrating data from both the machine and upper enterprise levels (Groba et al., 2007).

Resulting case study

Redundant, time-consuming, and manual maintenance can be avoided through an adequate information exchange between shopfloor environments (machine level) and ERP systems (Enterprise Resource Management), which is called vertical integration. The enhancement of information flow between the two different levels requires all significant information. Figure 2.4 depicts a schematic illustration of vertical integration (Groba et al., 2007).
In general, the machine's reliability is determined by proper health statements. For the machine operator on the shop floor, such a statement is critical in order to detect anomalous behaviour early. This statement with a one-dimensional source (source = machine level) can be paired with branch-level production and planning estimation. The ERP contains all the necessary data on production plans, orders, and people, as well as branch-specific data, allowing the statements to be enhanced and fine-tuned with the most up-to-date upper-level information. This new bidirectional ERP interface will facilitate appropriate (decision) reactions (Groba et al., 2007).
The vertical integration implies the connection between the machine and upper enterprise-level

benefits every part of the organization. This is in particular to the optimization of maintenance operations regarding their efficiency (Groba et al., 2007; Panfilov & Katona, 2018).

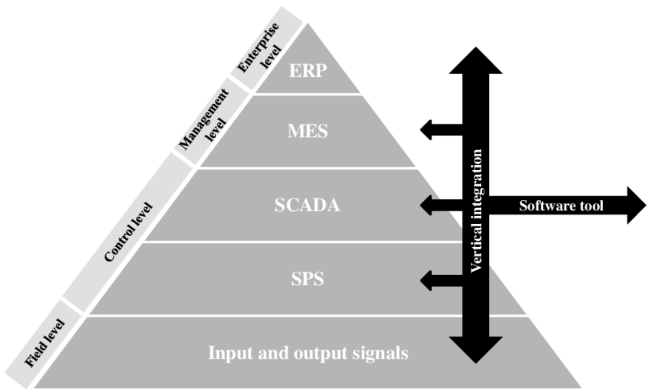


Figure 2. 4: The schematic representation of vertical integration within the automation pyramid. Information across non-intersecting layers can be combined to produce appropriate statements for better decision reaction (Thomas Vollmer & Robert Schmitt, 2015).

2.2.3.3. Problem- resulting case study 3: keeping the virtual asset up to date with the physical asset

Problem

During equipment modifications, mostly during weekends, there is a relatively short time frame in which the plant must be physically replaced and programmed. If the change cannot be made before the deadline, everything must be removed so that manufacturing can begin the following working day. Virtual assets like DT can help here by reducing the negative consequences of the abovementioned turnaround. Not only would this eliminate the need to define the planning's starting point, but it would also speed up the process by adopting virtual commissioning methods, which allow the software to be tested and optimized concurrently with hardware modifications. Furthermore, generic process optimizations can be tested first in a virtual mode (Zipper et al., 2018).

Unfortunately, the virtual asset and the physical plant will diverge as physical changes are made but not updated within the virtual asset. This is because maintenance and process optimizations cause physical and logical changes in a plant. Furthermore, a greater focus on customer needs to be combined with shorter innovation cycles causes changes in production equipment, resulting in an increasing number of variants for which production equipment must be adjusted or reconfigured. These changes are not registered and updated within the DT, so it is not anymore to the proximity of an identical twin to the physical plan. This leads to a DT, which could not serve the planning and commissioning of turnarounds in virtual mode (Zipper et al., 2018).

The challenge is to systematically converge the virtual asset (DT) and Physical asset (plant) by synchronizing the DT with the physical changing plant over the whole DT life cycle.

Resulting case study

A methodology provided by Zipper et al.is updating the model modification by employing monitoring. Continuous monitoring helps identify even very slow changes in the plant by continuously comparing simulated data with the real plant and helps confirm that the real plant and DT is still identical (Zipper et al., 2018).

This is approached by comparing the monitoring of the DT with factory measurements. Deviations from individual monitoring samples can be detected in this way and then pre-processed by classifying possible failures or deviations into several clusters (Zipper et al., 2018).

The preprocessed data is fed into an evaluation system to assist the modifier in updating the adjustment of the relevant monitoring models. The updated model should be tested to avoid manual labour inconsistencies in the overall model. After passing this check, the factory representation can be updated. This triggers a process to update individual monitoring models (Zipper et al., 2018).

2.2.3.4. Problem- resulting case study *4: Visualize the results in repair and replacement time periods*

Problem

Effective maintenance strategies extend lifetime and reduce maintenance costs. PdM is a promising strategy that detects asset failures, monitors degradation from historical data, and predicts the asset's future lifetime. However, it is unclear what and how the results from the PdM should be visualized concretely to improve the AM decision-making (Cheng et al., 2020).

Resulting case study

Cheng et al. simplify the result for the maintenance planning in two Remaining Useful Lifetimes (RUL) per asset. The first RUL estimates the period of when the asset should be *repaired*, for instance, a period over 7-14 weeks. This insight for the asset manager assists the planning of appropriate maintenance activity.
Moreover, the second RUL type, 'repair RUL', is provided for the same asset. This is applicable if a repair is not possible. An RUL replacement output could be:' Replace the asset over 14-20 weeks'. This assistance of repairing and replacing RUL per asset is beneficial as more assets are implemented in the PdM model and continuously update the asset manager about the overall plant health (Cheng et al., 2020)..

2.2.3.5. Problem- resulting case study *5: Use Remaining Useful Lifetime (RUL) as additional AM indicators*

Problem

Problem 5 is similar to problem four, as found in section 2.2.3.4.

Resulting case study

RUL is a standard that is generally applied in PdM literature. RUL estimates the expected failure time when it crosses the defined maintenance threshold (Kim et al., 2021).

2.2.3.6. Problem- resulting case study *6: data collection*

Problem

The required data for feeding the PdM model could come from various sources with different maturity levels. Conversely, which sources should be employed to feed the PdM and the reasoning for using these sources are not clear (Panfilov & Katona, 2018).

Resulting case study

Data can be gathered from a variety of sources for the PdM. The two primary data sources used for the PdM model for further predictive analytics and reusing it are SCADA and sensors. The sensors generate unclean, incomplete, and low-quality data, which must be cleansed and processed.
SCADA data is typically pre-filtered; thus, it requires minimal processing. The disadvantage of employing SCADA data is that it provides less flexibility and rigidity for computations because the expanded sensor data is not logged when distilling to the SCADA data (Kim et al., 2021; Panfilov & Katona, 2018).

2.2.3.7. Problem- resulting case study *7: saving centralized and decentralized*

Problem

The best PdM choices are made when suitable decisions are made at the right time, with the correct data at the right organizational decision level (Choubey et al., 2019). In this regard, a common practice is to send all data acquisition to a central data warehouse. The result of this strategy is centralized decision-making. However, this strategy goes against the idea of sending the right information at the right time from the correct location because all data is centralized while all the equipment are decentralized over an area. As a result, haltering the suitable choices and slowing down the decision speed (Choubey et al., 2019; Kim et al., 2021).

Resulting case study

As an alternative, distributed decision-making could be utilized. PdM benefits from distributed decision-making for the following reasons. Most decisions, such as comparing current data to an acceptable level or issuing alerts, necessitate immediate action. Waiting for a central place to make a decision is not cost-effective due to transmission time and costs (Choubey et al., 2019).

However, in some circumstances, centralized decision-making is necessary, but it should be limited to a bare minimum for effective and optimal data transfer. When the equipment fails to act as expected based on the features learned by the model placed at the local site, a decision must be made at the central level (Choubey et al., 2019; Kim et al., 2021).
As a result, a combination of localized and centralized decision support systems should be utilized. Here, the Local decision support will operate as the main server.

2.2.3.8. *Problem-* resulting case study *8: knowledge DB saving scenarios*

DeP 9: Save scenarios and known problem-solution pairs at the knowledge DB

Problem

Information storage must be efficient for the search to be completed in the shortest amount of time possible to deliver the best possible client experience. Commonly, similar problems reoccur that require the same solution. However, the solution is reproduced in an ad-hoc and iterative manner. It is meaningless to come up with new solutions to a problem that has already been solved successfully. An efficient way should be utilized.

Resulting case study

The attained knowledge from the PdM model and the pairing of problem to solutions (referred to as problem-knowledge pair) can be stored in the knowledge database, so when the next time a problem is discovered or reported, this knowledge base can be utilized first to find a solution, before generating a possible new solution to this problem (Choubey et al., 2019).

2.2.3.9. *Problem-* resulting case study *9: Operator training system (OTS)*

Problem

Integration of Industry 4.0 components and the advancement of digital technology can lead to more efficient and adaptable processes, allowing for the production of a wide range of high-quality products at lower costs and faster, giving a substantial competitive advantage. Cyber-Physical Production System (CPPS) deployment within Industry 4.0 has a significant contribution to many processes. Thanks to more relevant analysis and management. A CPPS must have a digital representation of the real system to make optimal judgments (Havard et al., 2019).

However, this form of CPPS, which is a DT, it is not always evident to operators and contractors how the digital twin can offer value and how business models can be built in the digital world. This is especially true for the contribution of DT to operator training (Bamberg Andreas, 2021).

Resulting case study

Operator training has broad applicability across the literature. Curl et al. suggest learning operators the ability to operate within exceptional situations in a safe virtual environment, like emergency shutdowns and equipment failures.

Also, operators could test their own operational theories within the operator training simulation within malfunctioning equipment or other undesired consequences.

Harvard et al. suggest educating principles, safety procedures and learning to assemble and disassemble parts when intervening in assets. These procedural maintenance skills within the DT are transferable to real assets without the need for gamifying the learning process (Havard et al., 2019).

Leingang et al. propose that operator training can benefit from building know-how on the plant behaviours. Here, normal and abnormal situations can be trained to minimize the operators' reaction times. These concepts should be associated with didactic when building training scenarios (Leingang et al., 2019).

In general, the contribution that DT can make for operator training is to learn to 'operate' in normal and abnormal situations, refresh safety procedures, and understand the assembling and dissembling of assets to minimize reaction time. Additionally, the DT provides a safe environment where undesired consequences are acceptable, and the operations and maintenance are taught in a didactically manner.

2.2.3.10. Concluding Fields within DT and PdM

In total 9 problem-resulting case studies are critically assessed. Every single is unique and so the core concepts are displayed in a tabulated format in Appendix C.2 . columns 1 to 3.

2.2.4. Fields within DG

As mentioned from section 1.5.1 the AM-field within the DG scientific literature was barely researched. Therefore, a bird's eye view of four major DG topics which has a broad potential applicability are provided in this section. This foundation will serve as an initial step on realizing data governance within RWS.

2.2.4.1. Data life cycle

The data life cycle management can be optimized, reduce the workload and increasing the performance using intelligent softwares, like HADOOP. This type of software is conveniently used for reduces and optimizes big data (Rahul & Banyal, 2020).

2.2.4.2. Data ownership contracts

Rarely can policy keep up with the development of new technologies, which is also no exception for data ownership. The struggling part of data ownership is who (the data farmer or the data landlord) is responsible for the data, which is an intangible asset. This led to no clearly defined laws covering data ownership (Ellixson & Griffin, 2016).
Instead of implementing laws and regulations for data ownership, the focus should be shifted to how data farmers collect and manage the data in a more transparent and trusted approach. This contributes to better control, access and benefits every stakeholder (Wiseman Associate Professor of Law et al., 2018).

2.2.4.3. Sensitive data sharing

Sharing data with multiple parties is a practice used increasingly, so the potential of misuse and data leakages is present. By using emerging technologies, these risks could be minimized, ensuring the security of the shared data. It is the obligation of owners and managers to maintain their own security when managing massive data (Alladi & Prasad, 2018). A viable way is using a data-sharing system that identifies the leaks using a knowledge base algorithm. This proactive security facilitates accountability of misusing parties like, for instance, an unauthorized user opening a classified map (Zhang et al., 2019).

2.2.4.4. Data stewardship

Defining new roles related to the five domains managing the data is crucial to successfully implementing data governance. Some of the roles corresponding to the data governance domains are mentioned in figure 2.5, and a whole data governance approach is built on defining roles and allocating responsibilities in table 2.1. An essential role for introducing data governance to the firm is the data steward. Hence, this section focuses on the crucial role within data governance.

Data stewardship is defined as a function between IT and business that is responsible for aligning business needs with IT for decision and operational support and executing the data governance program by entrusting them with data and metadata management. These functions are not owners of the data, rather ensuing the accuracy, accessibility, usability and keeping the data and metadata up to date. Their purpose is to use data as an asset to its fullest potential (Marco, 2016).
Data stewards consist of a team instead of an individual. Four data stewards' roles are found in almost in every organization (Marco, 2016):

Executive sponsor: A highly ranked, credible executive supporting and funding the data-related projects. He deals with political challenges, and therefore it is the most challenging type of steward to find.
Chief steward: This steward is responsible for managing the stewardship team. Therefore, he should be knowledgeable in the technical as the business stewardship for leading the following stewards: business stewards and technical stewards.
Also, he should be of senior level within the organization and have strong leadership skills as he will primarily manage the data governance projects and activities.
Business steward: This type of steward's responsibility is to make procedures and policies, create definitions to data, and search for business requirements that data can fulfil as an asset. The business steward is accountable to the chief steward and works closely together with technical stewards.
Technical steward: This steward is part of the IT department and focuses on the technical aspects of the data and metadata, data quality security requirements, giving definitions to technical data, and so on forth. The technical steward is accountable to the chief steward and works closely together with business stewards.

2.3. Conclusion of literature

Abstract concepts mentioned in the introductory chapter have broad and differing definitions resulting in ambiguity. These abstract concepts were asset management, predictive maintenance, digital twins, data governance, design-based research, and design principles. For this, the focus was to use the current scientific literature to define these abstract concepts for this thesis.
In line with the latter abstract concept, 'design principles', the found PdM and DT design principles were too generic and not specific for the AM. Therefore, 9 PdM & DT problem-resulting case studies were found that could have potential application within the AM context.
In addition, the AM-field within the DG scientific literature was barely researched. Therefore, a bird's eye view of four major DG topics with broad potential applicability is provided: data stewardship, sensitive data-sharing, data ownership, and data life cycle.
As a result of this literary background, the DeP's can be developed more rigorously.
The 'design principles' and 'design based research' final definitions are utilized to introduce the following chapter: the research methodology.

3. Research methodology

Chapter 2 briefly defined Design-Based Research (DBR) and Design principles (DeP) and the link between them. This chapter builds further by implementing it within the research design and the methods used for collecting and analyzing data. Firstly, the research design is elaborated in section 3.1. Subsequently, the two methods used for data collection are described in section 3.2. Finally, a series of four analysis methods are covered in section 3.3.

3.1. Research design: Overall approach

This research aims to identify design principles that overcome asset data problems during the implementation of DT, PdM and DG within the asset management context. These DeP's can provide broader applicable directions useful to academia and industry.

The research design for this thesis is built up of three main components. The first component is adopting a Design-Based Research (DBR) for producing DeP's. The DBR approach improves learning by developing the appropriate solutions to real-world problems systematically. This is executed following the three-cycle view of Hevner. The three cycles comprise of relevance[1], rigor[2] and design cycles[3] (Hevner, 2007). DBR is previously elaborated in section 2.2.1 (Herrington et al., 2009).

The second component is that the DBR is qualitatively approached. This approach is suited as it is explorative and descriptive, so minimal inference is exerted to the normal workflow. This is ideal for the work environment at RWS. The time horizon is cross-sectional due to the limited time period. (Uma Sekaran, 2018)

The last component for the research design is defining the types of design principles that will be produced:

- **Interview-derived Design Principles** (Int-DeP): Providing the empirically found principles at RWS.
- **Literature-derived Design Principles** (Lit-DeP): Providing the literature's current known principles.
- **Refined Design Principles** (Ref-DeP): Facilitating novel principles that are directly implementable at RWS by combing Int-DeP's and Lit-DeP's.
- **Data Governance Design Principles** (DG-DeP): Facilitating data governance principles relevant for RWS.

As a result, of using the three-cycle viewed DBR as the research methodology (1), which is qualitatively approached (2) to formulate four types of design principles (3), a schematic overall methodological approach is provided in figure 3.1. The following two sections will elaborate on every component in figure 3.1.

---

[1] The relevance cycle: providing the input requirements from the environment (people & organizations) in order to identify practices.

[2] The rigor cycle including the experience and expertise to guarantee that the contributions is not a routinely research based on well-known processes.

[3] The design cycle iterates between the findings of the rigor and relevance cycles in order to generate alternatives till a satisfactory design is established.

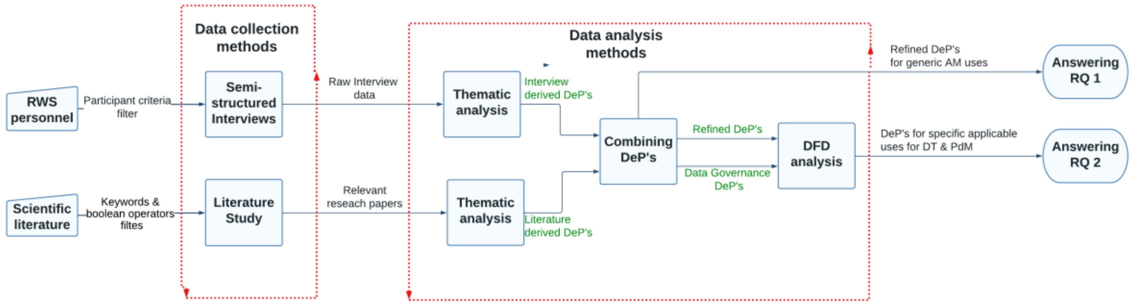


Figure 3. 1: The overall Design-Based Research (DBR), including the used research methods to answer the main research questions.

3.2. Methods of data collections

The input for the DBR is twofold, primary and secondary data. Both types of data are qualitative in nature and are utilized by the following two research collection methods:

- Semi-structured interviews
- Literature review

3.2.1. Semi-structured interviews

The semi-structured interview is a qualitative data collection method, and it is characterized by producing primary data because the researcher directly collects it. (Uma Sekaran, 2018)
The reason for using semi-structured interviews is to ask open-ended questions so that the interviewees can verbalize and share their own experiences and insights regarding the topic and allow the researcher to ask relevant follow-up questions. After the thematic analysis, the semi-structured interviews facilitate the formulation of interview derived DeP's.
The mediums used for data collection are Microsoft Teams (1) to plan and hold tele-meetings due to 'remote working' and a self-made questionnaire (2) which blends predetermining Interview Questions (IQ) to guide the semi-structured interviews. Also, the questionnaire enables a convenient data analysis. To build the questionnaire, operationalization from abstract concepts to IQ's and the grouping & order of the questionnaire are executed as these are important prerequisite steps.

3.2.1.1. Questionnaire development

Operationalization: A vital prerequisite for building the IQ's is to reduce abstract concepts to measurable and tangible variables that could be asked as an IQ and efficiently collected and analyzed. (Uma Sekaran, 2018) In this research, every concept is abstract: ''asset management, predictive maintenance, digital twins and data governance''. Operationalizations is executed in two steps. Firstly, a critical analysis of literature defines the abstract concepts, which are found in section 2.1. Subsequently, in appendix A.1, the literary defined concepts are further translated into IQ's.

Constructing the interview questionnaire: After operationalizing the abstract concepts, the questionnaire structure is determined. The questionnaire is divided up into three main sections as found in appendix A.2.:
The first section of questions regards DT and PdM for asset management and their developments in

their field. The second part covers data governance IQ's, and the last section rounds off the interview by closing questions like asking recommended interviewees, additional documentation, and questions from the interviewee towards the interviewer.
Importantly, section 1 of the questionnaire is divided into two parts and is asked depending on the interviewed focus group. Focus group 1 are interviewees working in the RWS lock complex division, which is the most developed regarding PM, DT and data governance. Here, best practices, advice, insights & demos are shared that can be reproduced at the SSB division instead of "reinventing the wheel". In other words, questionnaire section 1, part 1, is asked to focus group 1.
The other interviewees, focus group 2, working at the SSB division, are asked about the current stage of development and their future perspectives per SSB. The answers per SSB will be compared and analyzed. Therefore, questionnaire section 1, part 2, are asked to focus group 2.

3.2.1.2. Sampling size

The sample for semi-structured interviews was RWS personnel (individual = unit of analysis). The contact list was provided by a state trainee (rijkstrainee) who is involved with various data-driven technologies projects and holds a relevant network within RWS.
Also, during the interviews of the selected respondents' other eligible respondents were recommended to contribute to this research broadening the sample size.

3.2.1.3. Participant criteria

As a result of section 3.2.1.2, an initial list of eligible respondents who can grow was filtered. This respondent selection was achieved using the following enumerated criteria:

- **The respondent is 18 years or older:** This criterion set by the Human Research Ethics Commission for research using human subjects, i.e., interviews.
- **Experience level:** The respondent has a minimum of 1 year of relevant experience or knowledge about maintenance, data-driven technologies like PM, DT, or topics within data governance. This requirement ensures that the respondent is aware of emerging technologies or knows about the current need of asset management and has general knowledge about the associated terminology mentioned in section 2.1.
- **The respondent is filtered based on the interviewer and supervisors:** the vital stakeholder will select the participants regarding this report.
- **Focus groups:** The respondent works in one of the two focus groups within RWS: SSB or lock complexes. The lock complexes domain is the most developed regarding PM, DT and data governance in RWS. Here, best practices, advice, insights & demos are gathered that can be reproduced by the SSB instead of "reinventing the wheel".
  The SSB domain is the target for implementing emerging data-driven technologies to improve asset management. Here, questions regarding ongoing initiatives and the demand of asset management that PM, DT and data governance may fulfil will be interviewed

However, it could occur that the selected respondent does not meet one or more of these requirements during the interview. These interviews will be labelled as 'non-relevant interviews.

3.2.1.4. *Language of the interviews*

The language used for this research is English. However, the interviews are conducted in Dutch, the primary language used within RWS. Consequently, all validated interview notes were translated from Dutch to English. This translation ensures informational integrity by matching the sampling data language with the language used for theoretical concepts from literature.
If a non-Dutch speaker is interviewed, the communication will be in English, bypassing the translation step.

3.2.1.5. *Human Research ethics*

For every research involving humans as a subject for sampling data, the Human Research Ethics Committee (HREC) of the TUDelft must be approved. Unfortunately, the approval process was disrupted due to technical problems during the admission of the approval. Therefore, the TUDelft supervisors made an exemption for the semi-structured interview with RWS personnel in this thesis.

3.2.1.6. *Reliability and validity*

The following set of guidelines preserves the reliability and validity of the data collection.

1. **Validation of respondents:** The search and selection of the respondents are based on the 'participant criteria' list as found in section 3.2.1.3. The respondents should meet all the requirements before being used as sampling data.

2. **Invitation with related attachment:** In the invitation, agenda points were introduced with relevant information regarding the topics. This sent information inform prepares the interviewees and sets the interview's tone.
   Moreover, the interviewee can decline the invitation as they may not have the required expertise. This indicated that the researcher's preliminary investigation was incorrect.

3. **Structured interview approach**: All the respondents were provided with the same information, and the same questionnaire was used to analyze a valid sample.

4. **Confidentiality**: The interviews held via MS Teams were not recorded due to confidentiality. Instead, notes were made by the researcher. Afterwards, the notes were validated by the interviewee before using them as a data sample.

3.2.2. Systematic literature study

A literature study is a qualitative data collection method. The raw data are research papers and are labelled as secondary data because the researcher did not have an influence on the paper's collected data. (Uma Sekaran, 2018)
The mediums used for data collection are Google scholar and the TUDelft library as sources (1) for accessing research papers and books, and Mendeley Reference Manager (2) for managing and critically reviewing the literature in a structured way.

This literature search required two filter mechanisms: keywords and boolean operators.
The keywords that were used for the literature search are: 'asset management', 'predictive maintenance', 'digital twin', 'design principles', 'data governance', 'data governance design principles',

‘best practice’ and the combinations between them. However, the hits found were consistently over the millions, which was undesired.
Hence, boolean operators were used to only present papers which have predetermined exact phrases. This resulted in a significant decline of 51100 hits with the keywords ‘data’ and ‘governance’ combined with predictive maintenance compared to the 3400 hits using boolean operators with the keywords *‘data governance’* in combination with predictive maintenance. For the later case, only data governance literature was displayed and excluded all ‘governance’ and ‘data’ key terms using the Boolean operator.

As a result of using the two filters, the literature background in chapter 2 is compiled.

3.3. Methods of analysis

The two qualitative data collection methods provide input for analyzing the derived literature and interview derived interviews. This section will cover how the semi-structured interviews sample and literature review samples were analyzed to answer the research questions. In total, four analytical steps were utilized.

3.3.1. Relevance cycle: Thematic analysis of semi-structured interviews

After the semi-structured interview sampling, the verified notes were thematically analyzed to formulate Int-DeP’s and data governance problems. This empirical derivation from the environment producing practices is the relevance cycle from Hevner (Hevner, 2007).

five patterns were analyzed: type of Data-driven Project (DDP), encountered DT or PdM problems, the associated best practices, the SSB demand from the DT or PdM technology, and the DG problems. From the first four patterns, Int-DeP’s were produced, and from the latter four patterns, sub-RQ 1.1-1.4 are satisfied.

Afterwards, new themes relating to Int-DeP’s were explored using coding. The first step was axial coding. The DeP’s were compared and categorized into groups displaying new dimensions. Thereafter, selective coding was executed. The categorized groups are tied together, formulating one or two core categories helpful in determining the applicability for the DFD analysis in section 3.3.4.

3.3.2. Rigor cycle: Thematic analysis of literature study

The literature study across three scientific literature fields was also thematically analyzed. Problem-Resulting case study patterns within the PdM and DT literature were found, resulting in Lit-DeP’s production.
Moreover, the data governance literature contributed specific solutions to the problems found from semi-structured interviews with RWS -personnel. This inclusion of experience and expertise is the rigor cycle from Hevner (Hevner, 2007).

3.3.3. Design cycle: Combining Design Principles

The Lit-DeP and Int-DeP produced from the thematical analysis of the literature study, and semi-structured interviews are combined to a novel type of DeP, also called the refined DeP (Ref-DeP). These refined DeP’s will take the core insights of both Lit-DeP’s and int-DeP’s, which will be contextualized independently by summing up the principles, establishing a sequence of steps or confirming the found Int-DeP by the Lit-DeP and vice versa.

The DG-DeP’s were produced by firstly examining the encountered RWS problems from the semi-structured interviews. For these problems, a solution is provided from the up-to-date data governance literature. With the set of Ref-DeP’s and DG-DeP’s, the main RQ1 is satisfied.

3.3.4. DFD-analysis

The purpose of the DFD analysis is to test the principles in practice. This is achieved by comprising the DFD analysis into three steps: building the DFD, implementing DeP’s and identifying opportunities and bottlenecks.
Firstly the two DFD’s are built. The DF for the DT (DFD-DT) is developed by collecting information based on the currently operating DT at the SSB1 from the DT architect at RWS. Conversely, DFD-PdM is built 100% from the Ref-DeP’s as no PdM model is currently operated at RWS from which the DFD-PdM could be inspired. The program used to build the DFD is Lucidchart which is open source.
Secondly, the applicable and relevant Ref-DeP and the DG-DeP’s were implemented into the DFD’s. This means that not all generated DeP’s are used in the DFD. With this, sub-RQ 2.1 and 2.2 is answered.
Lastly, the opportunities and bottlenecks are analyzed. These are found between the interacting DeP’s providing a deeper layer of insights that would not be independently found as a standalone DeP’s
With the set of opportunities and bottlenecks, sub-RQ 2.3 is satisfied, and with this, RQ2 is answered.

3.4. Summarized research methodology

The research design for this thesis is built up of three main components. A three-cycle viewed DBR as the research methodology (1), an approach that is qualitatively in nature(2) and lastly, the formulation of four types of design principles (3). These four are Interview-derived Design Principles (Int-DeP), Literature-derived Design Principles (Lit-DeP), Refined Design Principles and Data Governance Design Principles (DG-DeP).
The data will be collected using semi-structured interviews and a systematic literature study. For the semi-structured interviews, a questionnaire is built to guide the vast range of topics during interviews. Subsequently, a series of data analysis steps occur. Firstly, by thematically analyzing the semi-structured interview data to produce Int-DeP’s. Secondly, by thematically analyzing the literature study produce Lit-DeP’s. After that, these results are combined to produce Ref-DeP’s and DG-DeP’s satisfying main RQ1. Finally, the set of Ref-DeP’s and DG-DeP’s, are analyzed within a DFD model to identify opportunities and bottlenecks when implementing DT, PdM and DG. With this main RQ2 will be satisfied.
This research methodology chapter provides the overall underlying strategy for realizing the desired results in the following three' results chapters'.

4. Results – RWS semi-structured interview analysis

This chapter produces interviews derived DeP's from the executed semi-structured interviews at RWS. First, a quantitative analysis is provided to argue about the interview's sampling size and considerations on the sampling confinements. Afterwards, the defined terms (PdM, DT, AM, DG) from the literature background chapter were used to clearer analyze the data-driven projects (DDP), encountered problems, best practices, and SSB demand. As a result of combining the beforementioned enumeration, Interview-derived DeP's are produced.
In addition, an approach is attempted to formulate the current data governance problems at RWS.

4.1. General quantitative analysis semi-structured interviews

In total, 42 interviews were held with RWS employees resulting in 6 found DDP across lock complexes and SSB. Table 4.1 displays the general analysis of this vast number of semi-structured interviews. The considerations of the sampling size and the confinements are made as follows:

Firstly, the 33 effective interviews indicate the generalizability of this research. This means that the findings are more reliable, easily transferable, and comparable to other similar socio-technological studies.
The non-relevant interviews, nine interviews of the 42, did not contribute to this research. These were interviews focusing on general working methods within RWS and thus did not answer the IQ's from appendix A.2.
Another note is the number of interviewees at the lock complexes. This is significantly smaller than at the SSB. This is because the lock complex candidates were in the finalizing phase of the respective DDP's. For this reason, it induced constructive, clear, and smooth interviews, giving sound guidance, answering questions extensively and providing demos.
On the other hand, the number of SSB interviews was expansive for two reasons. Investigating for DDP similar to the lock complex and inventory the SSB demand for DT and PdM. The latter one required, called brainstorming sessions, were interviews where multiple experts were invited per SSB, justifying the high number of interviews.
Lastly, only three from the 6 SSB were sampled for this research. The reasoning is that many SSB function similarly. Therefore, three representative SSB were chosen based on the operating philosophies. SSB1 follows a multi-disciplinary team managing the closure, SSB2 closes the walls automatically by the operating system, and SSB3 is the oldest operated with outdated manual working methods to close the SSB.

Attribute	Number
Total interviews	42
Non-relevant	9
Effective interviews	33
Data-driven projects (DDP)	6
Lock complex interviewees	8
SSB interviewees	25
SSB	3 out of 6

Table 4. 1: Quantitative analysis of the performed semi-structured interviews

4.2. Descriptive analysis: Describing DDP

This section will give a brief description of the DDP with the associated encountered problems and best practices which will be the foundation of producing Interview derived DeP's. The naming of the DDP's is given general labels due to the confidentially. With the identification of the encountered problems and best practices sub-RQ 1.1 and 1.3 are satisfied.

4.2.1. Lock complex division-specific projects

Three DDP have been identified from the lock complex: A Condition-Based Maintenance (CBM) project, a DT project, and a general systematic plan (4-phase plan) to make assets more data-driven. Every DDP highlights the encountered problems and generates corresponding best practices in tabulated form.

4.2.1.1. DDP 1: generic 4-phase plan

The vision of the lock complex division is to move from traditional time-based maintenance to a data-driven form of asset management. This should be a systematically approach which is less dependent on ad-hoc working methods.
However, there are two encountered problems. The first lacks a step-by-step actionable plan that is generic and applicable for every asset within RWS.
The second is the difficulty to track multiple projects concurrently and communicating effectively. These innovative, ongoing projects are mostly not documented and unknown for other divisions. The lack of transparency could lead to two or more simultaneous redundant projects. In the ideal scenario a project should share it in-field practices so it can prevent the 'reinventing of the wheel' in different divisions.

As a result, the 4-phase plan was established with two main best practices. It outlines how every asset should generally follow four distinct consecutive steps. In addition, in-between the four steps, a go/no-go gate continuously evaluates the progress for its relevance, time and investment of the ongoing project.
The second is the periodically organized Community of Practice (CoP) to share, analyze, and transform learning experiences in implementing data-driven technologies into tools, specific instructions, training and prevent redundant projects. Table 4.2 displays all the encountered problems with the associated best practices.

Int-DeP #	Encountered problems	Best practices
1	No step-by-step actionable plan that is generic and applicable for every asset to make it data-driven within RWS.	A generic 4 phase plan with in-between go/no-go that is applicable for every asset.
2	Difficult to track multiple projects concurrently and communicate effectively to evaluate redundancy of projects and share in-field practices.	Recurring organized CoP to share, analyze, and transform learning experiences in implementing data-driven technologies into tools, specific instructions, training and preventing redundant projects.

Table 4. 2: Table of the encountered problems and the associated best practices found from interviews of DDP 1

4.2.1.2. DDP 2: DT Pump station

Currently, the administration of malfunctioning assets at the pump stations is reported manually. The process is by 'copy pasting' the relevant document to the associated windows folder. This method is prone to human errors because muddling through by endlessly opening windows folders searching for a questionable asset tag number will lead to fault registration at the wrong file. Therefore, incomplete insights into the asset status indirectly led to increased maintenance. Therefore, incomplete insights into the asset status will indirectly lead to increased maintenance.

The best practice for the mentioned problem is done by implementing a DT displaying the real machine hall full of assets (and many components per asset) that are clickable to the directories of the existing AM software. This user-friendly interface minimizes the human incorrect asset malfunction registration. An expansion of this project is to make the virtual asset clickable from the DT for every AM software to their operational and strategic objectives. Figure 4.1. displays the information handling from the physical asset to the AM software by comparing the current manual situation and the ideal DT intermediating situation.
A different problem occurs to keep the virtual model (DT) up to date. When physical changes are made, i.e., changing a pump, the DT should be updated to prevent information discrepancies between the physical (pump station) and the digital asset (DT pump station). The prescribed best practices are to integrate the DT update in the management of change (MoC) procedures. However, this is not implemented and tested on its effectiveness. Table 4.2 displays all the encountered problems with the associated best practices.

Int-DeP #	Encountered problems	Best practices
3	The human error of Incorrect asset fault registration influences the maintenance schedule negatively.	Use the DT as the user-friendly interface for translating physical asset information to the existing AM software by clickable directories per asset(component).
4	Keeping the virtual model up to date	Implementing DT updates during the MoC procedures.

Table 4. 3: Table of the encountered problems and the associated best practices found from interviews of DDP 2

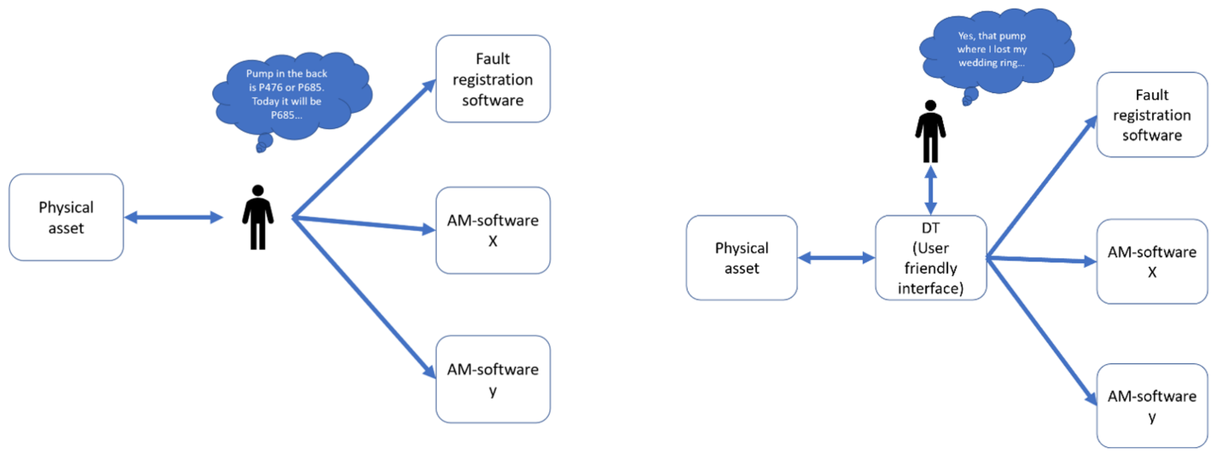


Figure 4. 1: Current situation (left): The user administrates every decision between the physical asset and the vast array of AM software compared to the ideal situation (right) in which the user-friendly DT facilitates improved user experience to meet the strategic, operational goal.

4.2.1.3. DDP 3: Data visualization hydraulics

Across RWS, the current dominant maintenance strategy is preventive maintenance which is characterized by fixed maintenance activities over the asset life cycle. This method is robust to forecasting maintenance budgets. However, it additionally results in over-maintenance and, in worse situations, reactive maintenance that is usually 3x more costly. More characteristics of preventive maintenance are found in section 2.1.2.
To improve the decision-making of AM, a higher degree maintenance strategy should be employed. For this, four best practices are found which facilitated this DDP. Table 4.4 displays all four encountered problems with the associated best practices.

The first is the carried-out best practice of implementing Condition Based Monitoring (CBM) instead of the ideal PdM. The reasoning is that CBM is more practical as it does not require the computationally intensive models and processing power as PdM, but it still has significant benefits like monitoring the historical and current status of the asset. Therefore, implementing the CBM before considering the PdM strategy is the first step forward to data-driven asset management.
The following best practice is to visualize the asset behaviour as pressure, temperature, and other computed AM parameters. These graphical representations are projected in Power BI to provide insights to the decision-makers for adjusting the maintenance activities. Before the implementation, it was unclear about the historical and current status of the asset and which parameters were important.
The next best practice is storing all the full and complete data except for data older than one year. The data older than one year will be aggregated and denoted as historic. The encountered problem is that much data is saved, but not much information is relevant. This is a waste of using a massive data center, slowing down the model. This 1-year demarcation best practice makes storing more manageable.
The last best practice is that the data sources are only used from the raw sensor data, and/or reusing the preprocessed data from SCADA as computing again wastes processing power. The associated

problem was that it was unclear which sources the data could be collected. This limitation to two sources will provide the best result with minimal effort.

Combining these four best practices avoids significant maintenance operations that were not found before the DDP3 implementation. For instance, a situation occurred at a pumping station where erroneous behaviour could be observed on the Power BI dashboard and compared with historical data. A (data-driven) decision was made to send a diver for underwater investigation. Surprisingly, large basalt blocks were found in front of the pump's entrance. The problem was solved easily by removing the basalt block. This would not be possible before the DDP3 implementation leading to the block being sucked into the pump and causing extensive damage.

Int-DeP #	Encountered problems	Best practices
5	No insights into the historical and current status of the assets.	Visualize the asset behaviour as pressure, temperature, and other computed AM parameters in PowerBI.
6	Unclear from which sources the raw data can be extracted.	Use raw sensor data and preprocessed data from SCADA.
7	Storing all the data but not much information is relevant leading to slowing down of the model.	Data older than one year is aggregated to the desired size and labelled as historic. Data < 1 year stays full and complete.
8	Transitioning to PdM is a giant step with many complications and uncertainties.	Implement CBM before considering the PdM strategy.

Table 4. 4: Table of the encountered problems and the associated best practices found from interviews of DDP 3

4.2.2. SSB division-specific projects

All 6 SSB's follow the rules and regulations enforced by the Water Law (Overheid.nl, 2022). The water Law encompassed all the related information and documentation for flooding in the Netherlands. The leading parameter in the Water Law is the 'SSB fail chance estimation', which states that every citizen has the same risk of death from flooding, i.e. 1 in every 100 closure may fail.
The challenge for the projects mentioning in the following sections is to keep the fail chance within the prescribed boundaries even during major hardware/software change, uncertainties and other disturbances.

4.2.2.1. DDP 4: DT SSB 1

Every 15 years, the Operating System (OS) of the SSB's must be changed. These consist of all Programmable Logics Controllers (PLC) and PLC software. There are two encountered problems associated with this significant recurring maintenance activity.

- The OS redesign from scratch is associated with uncertainties. This undesired occurrence is because teething problems like glitches, errors, and incorrect actions happen during new OS delivery. Hence, these software uncertainties are accounted for within the fail chance estimations for the specific SSB leading to undesired deviations which are not in line with the Water Law.

- The tacit knowledge for understanding the logics disappears every 15 years leading to redesigning the wheel (operating system) over again. The tacit knowledge is essential for building faster OS due to the conceptualization of the PLC functioning. However, not actively using the knowledge on a daily or weekly basis and the outflow of knowledge workers over a period of 2 decades negatively influences the capturing of this prevalent knowledge.

Thus, In connection to the mentioned problems, a DT is utilized for two purposes: capturing the tacit knowledge (1) and validating the new OS (2).
The new OS validation happens by first interacting with the DT, which has the same PLC functioning as the physical SSB. Various scenarios can be loaded from the database to test the OS for the teething problems. This virtual testing environment is superior to real testing as it also verifies the OS for unhappy flows. The unhappy flows in real testing are dependent on external factors (like weather and broken pumps), which are not easily reproducible. An additional side feature of validating the new OS is the speed-up function. In the actual scenario, SSB closure may take hours and requires while the DT 4 closures could be tested within 15 minutes.

The other purpose of the DT is to capture the tacit knowledge which is lost over the 15-year period. The DT holds the identical PLC functioning, and therefore the tacit knowledge can be preserved in a digital asset. The only setback is updating the virtual asset as the physical asset changes. This is a similar situation as mentioned in DDP 3 in section 4.2.1.3. Table 4.5 displays all the encountered problems with the associated best practices.

Int-DeP #	Encountered problems	Best practices
9	The OS redesign from scratch is associated with uncertainties leading to undesired deviating fail chance estimations for the specific SSB.	Building a virtual environment (DT) where the OS can be validated in happy and unhappy flow with speed up functions before commissioning it to the real SSB
10	The tacit knowledge for understanding the logics disappears every 15 years leading to tedious redesigning of the wheel (OS) over again.	Capturing the tacit knowledge and preserving it within the DT is associated with an updating strategy.

Table 4. 5: Table of the encountered problems and the associated best practices found from interviews of DDP 4

4.2.2.2. DDP 5: DT SSB2-A

SSB 2 consists of 2 operating systems A and B, both mutually exclusive and exhaustive. SSB2A's OS primary function is to control everything except for the walls (SSB2B OS primary function).
The encountered problems and best practices from DDP 4 are similar to this DDP. The only additional problem is that the number of closures of SSB2 is yearly the lowest (1.6 times per year) and the most advanced and autonomous system from all the SSB. As a result, inexperienced operators understand the closure theoretically but have minimal infield experience. This problem can be solved by implementing virtual training sessions using the same happy and unhappy scenarios for testing the SSB OS. The additional part of the training is the didactical factor. By utilizing the virtual training sessions, the

operators can become more proficient and confident in performing this complex closing procedure. Table 4.6 displays all the encountered problems with the associated best practices.

Int-DeP #	Encountered problems	Best practices
11	Operators have too little practical experience with closing the physical SSB and, therefore, undertrained.	Facilitate a simulator (virtual environment) where operators can be trained in a didactic manner in happy and unhappy flow scenarios.

Table 4. 6: Table of the encountered problems and the associated best practices found from interviews of DDP 5

4.2.2.3. DDP 6: DT SBB2-B

SSB-2B operates the walls combined with SSB-2A OS controlling the rest of SSB2. The encountered problems and best practices of DDP 6 are similar to DDP 4 & 5.
A DDP 6 specific problem, originating from a pure IT technical perspective, is that the DT's are produced in an ad hoc manner elongating the deployment every time significantly.
Hence, as a best practice, the experience from the previous 2 DT production is used to build generic configurable components to facilitate a shorter DT production delivery. The DDP 6 can build the first solid foundations of a set of configurable components for the future 2 DT projects. Table 4.7 displays the encountered problem with the associated best practice.
This project is still in the specification phase and thus still not executed. Therefore, no more information is found regarding DDP 6.

Int-DeP #	Encountered problems	Best practices
12	Designing the DT takes an enormous amount of time to commission fully.	Build/use generic configurable components based on the previous DT projects.

Table 4. 7: Table of the encountered problems and the associated best practices found from interviews of DDP 6

4.3. Descriptive analysis: SSB demand

The enumerated DDP from section 4.2. proposes ingenious data-driven ways to solve their AM problems. However, it is also essential to take the priorities and relevancy of SSB into account. Hence, brainstorm sessions were held to propose a number of DDP with the associated problem-best practice (DeP #) to various experts at 3 of the 6 SSB. The goal was to identify the attitude towards the implementation of new technologies which will transform the current AM working methods. Appendix B.1. shows which type of RWS-experts were invited to the brainstorming sessions. With the identification of the SSB demand for DT and PdM implementation sub-RQ 1.2 is satisfied.

4.3.1. SSB 1

The SSB 1 brainstorm session experts were interested in the DT related DDP's. DDP2: the interest was towards integrating the current AM software using the DT as the user interface (DeP #3). Also, keeping the virtual asset (DT) up to date (DeP#4) as of high interest indicates the importance of this development.

DDP5: Interest was towards the training programs of SSB 2 (DeP #11). Due to the outflow of experienced knowledge workers from the small dynamic team. Additional remarks were made to use next to the operational purposes also seems to be for learning maintenance activities where mistakes could be made (unhappy flow).

4.3.2. SSB 2

In contrast to section 4.3.1. the brainstorm session at SSB2 focused more on PdM's insights than on DT. The interest was particular toward DeP# 7 and 8 originating from DDP 3.

DeP # 7: Currently, much data is stored. However, the culprit can be that it will slow down the models.

DeP # 8: Currently, 2 HBO and 1 PdEng are exploring the possibilities of using the sensor data to improve maintenance activities. The of what is possible with the data. The first step is the go-to CBM and afterwards PdM.

4.3.3. SSB 3

SSB3 is one of the oldest SSB and, therefore, strongly dependent on the old working methods. The mindset was positive towards DT and PdM. Unfortunately, the demand was that the other RWS divisions should provide a total solution for the DT and PdM implementation before SSB3 considers providing insights into the asset. This is associated with the old ad-hoc working methods used up all the interviewees' time, leading to the delegation of innovation projects. In other words, SSB3 want to be unburdened entirely of the initial phase of new technologies implementation.
As a result, no DDP or one of the associated DeP was of interest to SSB3.

4.3.4. Contrasting views of SSB-demand

When the three SSB brainstorm sessions were compared, two opposing viewpoints emerged. The first is the SSB's openness to embrace new technology. SSB 1 and 2 are generally receptive to new data-driven technologies that can help them better manage their assets. SSB 3, on the other hand, is undoubtedly intrigued but has been less proactive in taking concrete actions. This finding assumes that there are differences between the SSB regarding the openness for embracing new technologies.
The following reasonings for SSB3's reserved attitude could be that traditional working methods at this SSB are time-consuming, leaving less time for embarking on new technology. It is also possible that the assets that have been in operation for decades are quite predictable, and thus there isn't a strong enough incentive to switch to data-driven asset management utilizing asset failure prediction or digital twin visualization.

Secondly, the alignment between asset management priorities and the respected data-driven technology. It turns out that every SSB is different with regards to its interest in the type of technology. SSB1 was interested in DT, SSB2 was interested in PdM, and SSB3 is still inconclusive. This varying interest originates primarily from the SSB's own set of priorities by which the technology can fulfil its needs. For instance, all equipments at SSB1 are still clustering at the start of the bathtub curve, which makes investments for PdM unfeasible. Another explanation could be knowledge transparency. The SSB2 is unaware of the full applicability of DT for asset management and therefore showed no interest.

These contrasting viewpoints unknowingly construct invisible (almost) impermeable walls, between different layers in the RWS organization, which prevents knowledge spillovers, departments maturing at different rates and impeding the understanding and communication between them.

4.4. Development of interview derived DeP's (Int-DeP)

Combining the insights of the encountered problems (section 4.2), best practices (section 4.2) and SSB demand (section 4.3) resulted in the formulation of 12 Interview-derived DeP (Int-DeP's). These 12 int-DeP's are shown in table C.1 in the appendix.
In general, the int-DeP's are similar to the best practice but are written in a prescriptive manner, solving a specific problem. The SSB demand indicates the urgency and relevance of the DeP and adds minor adjustments to the int-DeP formulation.

Further thematical analysis revealed two categories within the 12 formulated int-Dep's. These were:

Long term vs short term relevance:

It was noted from the SBB demand analysis that there were some int-DeP direct applicable and relevant for the SSB, while others were labelled as too generic, strategic and difficult to apply directly. Therefore, a new category emerged which groups the int-DeP in the short-term or long-term DeP's.
The long term was more strategic from nature and applicable for high-level DM to prevent future coming managerial problems. For instance, DeP 2, which organizes CoP, keeps track of the DDP projects and prevent multiple projects which could be redundant.
The short-term DeP's were directly applicable that solved acute problems faced during the implementation of DT and PdM.

DT/PdM/Both Relevance:

The derivation of int- DeP's is from a DT project or PdM project. Fortunately, due to the generic methodology of formulating design principles, some DeP's are applicable for both types of projects. For example, DeP 12 prescribed that DT production can be faster and more affordable by using generic configurable components. This principle is also applicable to producing PdM software.

4.5. Current status of data governance

Data governance term is not known within de RWS; therefore, the IQ's were tailored to ask about certain aspects of data governance that are more related for the RWS-interviewees, as shown in appendix A.1. With the results from this section, sub-RQ1.4 is satisfied.

The results from the interviews were overall poor as expected, indicating that the added value of data governance is unclear, unexplored and require more attention. A number of concerns were identified regarding the data life cycle, ownership of the data, and data access by third parties, as shown in table 4.10 below.

In specific, a major contrasting view was regarding the data access by third parties. The lock complexes performed many DDP's in collaboration with external maintenance parties. These provided their key resources and experience in harvesting and analyzing the asset data, which positively influenced the projects in many ways. On the other side, the SSB division is reserved and hesitant of sharing data because they are concerned that terrorists may gain access to sensitive information. The disadvantage of denying external parties access to data is that the transition to data-driven asset management will be greatly delayed by fostering the competencies that these parties already possess.

As a last point, the contrasting viewpoints, including section 4.3.4, are caused by unknowingly constructing invisible impermeable walls that prevent knowledge spillovers, resulting in departments maturing at different rates and worsening the understanding and communication between them.

problem	Potential demand
De data lifetime cycle is undetermined; therefore, all data is saved for an infinite time period till a decision is made.	An efficient way to store the data.
All the data ownership is at a third-party company, which leads to lower control on the asset as data.	Clarification of implementation of legislation paragraph regarding data ownership in the contract
Due to the SSB information being sensitive information for terrorism, third party collaborations are censed	Some form of data access should be allowed as some domain knowledge is RWS does not possess.
Unclear roles for integrating data-driven technologies for the AM	data scientist, engineers etc.

Table 4. 8: Interview derived problems and potential demand regarding the unclear role of data governance.

4.6. Conclusion semi-structured interview analysis

Chapter 4 started with the input of the semi-structured interviews data. These were qualitatively and quantitatively analyzed. There are three main qualitative results. Firstly, 33 effective interviews at RWS were held across the lock complex (8 interviews) and SSB (25 interviews) divisions. Secondly, 6 Data-driven Projects (DDP) were identified. Lastly, 3 of the 6 SSB were included in this thesis as the remaining are redundant because these SSB's function similarly to included SSB's.
Quantitatively, the 6 DDP's and the SSB demand from PdM and DT technologies were combined, resulting in a set of Int-DeP's. These Int-DeP's were further categorized based on 'applicable for DT of PdM or both' and 'long-term or short-term' relevance.
In addition, from the interviews, the current DG status described that DG problems were encountered for which no solution is currently found. As a result of this chapter, 12 Int-DeP were formulated, and 4 DG problems were found.
The interviews results revealed contrasting views on (1) the SSB openness for new technology, (2) the alignment between asset management priorities and the data-driven technology, and (3) data access by third parties. These contrasting viewpoints unknowingly construct invisible (almost) impermeable walls, between different layers in the RWS organization, which prevents knowledge spillovers, departments maturing at different rates and impeding the understanding and communication between them.

The Int-DeP's from this chapter are further used in the following chapter by combining it with the Literature-derived DeP's which produces the final set of DeP's.

5. Results – Development of ref-DeP & DG-DeP's

From the previous chapters, 9 problem-resulting case studies, 4 fields within data governance literature, 12 interview-derived DeP's and 4 interview derived data governance problems were collected and developed.
This chapter will transform the 9 problem-resulting case studies into 9 usable literature-derived DeP's and then combine the interview and literature-derived DeP's into a set of refined DeP's. Moreover, the data governance problems are coupled with the associated fields within the data governance literature to form a set of data governance design principles.
As a result, the developed refined and data governance design principles answer the main research question 1: ' What are the design principles for asset management?'

5.1. Literature-derived DeP

For the first step of developing the Ref-DeP's, the Lit-DeP's need to be transformed from the 9 problem-resulting case study from section 2.2.3. These 9 are focused on the DT and PdM implementation in the AM.
From transforming the problem-resulting case study to a Lit-DeP's required minimal effort because of the fact that the literature is a secondary data source which means it is already critically analyzed. The only computation needed was rephrasing the solutions in an actionable sentence.
As a result, a set of 9 Lit-DeP is developed. These 9 Lit-DeP with the associated problem-resulting case study are found in appendix C.2.

5.2. Refined DeP

The second step and last step for developing the Ref-DeP's is by combining the Lit-DeP (table C.2) and Int-DeP's (table C.1.) thematically. As a result, a set of 12 Ref-DeP's emerged. A tabulated form displays the 12 Ref-DeP's with the associated Lit-DeP's and Int-DeP's appendix table C.3.
Interestingly, during the Ref-DeP's development, it was noticed that four Int-DeP's (int-DeP's 1,8,10,12) had no associated Lit-DeP. Therefore, these Ref-DeP's were similarly formulated as the corresponding Int-DeP.
The remaining ref-DeP's formulations are distinct. Therefore, for every Ref-DeP, an explanation is provided below .

Refined DeP 1:

From the literature, no Lit-DeP is derived. For this reason, refined DeP will exactly be int-DeP 1.

The refined DeP: Follow the generic 4 phase plan for every asset, which guides, evaluates, and eliminates projects.

Refined DeP 2:

Lit-DeP 2 suggests that it should take five various knowledge processes into account when managing knowledge. From Int-DeP, two of the five are clearly stated by organizing periodic CoP's (knowledge process 1 = knowledge sharing) across division borders (knowledge process 2= boundary crossing) to manage the knowledge development.
This indicates that there is still room for optimizing the knowledge management by incorporating the

remaining knowledge processes: knowledge generation, collective project learning and knowledge integration. Consequently, by combining Int-DeP 2 and Lit-DeP 1 in the proposed manner, the refined DeP 2 states as follows:

Refined DeP 2: Organize periodic CoP to keep track of DT and PdM developments across divisions/SSB borders and optimize the knowledge management further by incorporating remaining knowledge processes: knowledge generation(1), knowledge integration (2) and collective project learning (3).

Refined DeP 3:

Int-DeP 3 proposes to enhance DT's by implementing other AM software's to facilitate a broader range of purposes. Lit-DeP 2 intensifies one particular purpose: enhancing machine reliability by including ERP data for more correct asset health statements, which improves decision reaction. Consequently, by combining Int-DeP 3 and Lit-DeP 2 in the proposed manner, the refined DeP 3 states as follows:

Refined DeP 3: Use the DT as a user-friendly interface by making it compatible with the currently used AM software through directories, so human errors are reduced and integrate adequate information exchange between the shop floor and AM software to enhance the decision reactions.

Refined DeP 4:

The challenge is keeping the digital asset up-to-date while the physical asset changes and is not documented. Int-DeP 4 recommends first making the changes in the virtual model before making the change. Lit-DeP 3 proposes an automatic correcting step that continuously compares the simulated data with the real plant to confirm that the real plant and DT are still identical. With these two steps, the digital asset can be kept up to date. As a result, by combining Int-DeP 4 and Lit-DeP 3 in the proposed manner, the refined DeP 4 states as follows:

Refined DeP 4: Test the change first virtually in the DT and update, then make the physical change and finally, complete it with documenting in the MoC procedure and concurrently continuous monitor by comparing simulated data with the real plant to confirm that the real plant and DT is still identical.

Refined DeP 5:

The data visualization currently displays the pressure, temperature and computed AM parameters like energy consumption to facilitate better AM decision-making. From Literature, an additional, computed parameter is using RUL's (Lit-DeP 5). This can predict when the asset needs maintenance. Additionally, Lit-DeP 4 proposes to use two types of RUL. One for the repairing and the other for replacement.
As a result, by combining Int-DeP 5, Lit-DeP 4 and Lit-DeP 5 by the proposed manner, the refined DeP 5 states as follows:

The refined DeP: Visualize the sensor data in temperature, pressure, RUL repair, RUL replacement, and other computer AM parameters to provide insight to facilitate maintenance planning.

Refined DeP 6:

The Int-DeP of collecting raw data and preprocessed data from SCADA is similar to the Lit-DeP. The only addition from the Lit-DeP is that the data should be reused. A way to reuse is found in the following in

refined DeP 7.
A condition for using refined DeP 6 is that every type of sensor data can be visualized. If, for instance, SCADA is not logging data for the dedicated asset and the sensor does is analogue, this DeP is then not applicable.
As a result, by combining Int-DeP 6 and Lit-DeP 6 in the proposed manner, the refined DeP 6 states as follows:

The refined DeP: Collect raw data at the source and the preprocessed data from SCADA and then reuse it

Refined DeP 7:

The Interview derived DeP proposes that the data storage should be split up into 2. Data that < 1 year old remains complete and untouched, and data older than one year will be aggerated as this data will be piled up throughout years and decades. From the literature derived DeP, the data storage strategy expands this by proposing data computations at two different locations based on complexity.
As a result, by combining Int-DeP 7 and Lit-DeP 7 in the proposed manner, the refined DeP 7 states as follows:

The refined DeP: Store data based on the data life cycle and computed data based on complexity. Data> 1 year goes to historical data, and easily computable data are solved by the decentralized level and harder computable data at the centralized level. This ensures the maximum effective DM and minimizes lag time.

Refined DeP 8:

From the literature, no Lit-DeP is derived. For this reason, refined DeP 8 will exactly be int-DeP 8.

The refined DeP: Implement first CBM, then transition to PdM strategy.

Refined DeP 9:

Uncertainties arise regarding the correct interaction between the new operating system(OS) and SSB, which leads to an undesired increased fail chance estimation. From the interviews, a solution is provided by using s database, which stores various happy and unhappy flow scenarios to test the new OS and cure the OS teething problems. This idea can be expanded for normal operations where similar situations arise repeatedly. For this, the same knowledge database can save known problem-solution pairs for advising known solutions for assisting the asset manager and operators when a problem arises.
As a result, by combining Int-DeP 9 and Lit-DeP 8 in the proposed manner, the refined DeP 9 states as follows:

The refined DeP: Store all the happy, unhappy flow and problem-solution pairs at the knowledge DB for reuse to prevent reinventing the wheel of solutions.

Refined DeP 10:

From the literature, no Lit-DeP is derived. For this reason, refined DeP 10 will exactly be int-DeP 10.

The refined DeP: Capture tacit knowledge and preserving within the DT.

Refined DeP 11:

Operators have too little practical experience closing the physical SSB and, therefore, undertrained.

The interviews facilitate training programs for operators within the DT environment. The literature defines normal and abnormal situations, Safety procedures, and assembling and disassembling assets concretely. By this, operators understand operations and maintenance didactically to minimize reaction time.
As a result, by combining Int-DeP 11 and Lit-DeP 9 in the proposed manner, the refined DeP 9 states as follows:

The refined DeP: Facilitate operators and maintenance to learn 'to operate' in the existing DT environments in various SSB behaviours, which are: Normal and abnormal situations, Safety procedures and assembling and disassembling assets. This minimizes the operations and maintenance reaction time.

Refined DeP 12:

From the literature, no Lit-DeP is derived. For this reason, refined DeP 12 will exactly be int-DeP 12.

The refined DeP: Identify generic configurable components to make DT production more affordable and faster deployed.

5.3. Academic relevant Int-DeP's

During thematically combining the Lit-DeP's with Ref-DeP's to develop Ref-DeP's, it was noticed that four Int-DeP's had no associated Lit-DeP. Therefore, these Ref-DeP's were similarly formulated as the corresponding Int-DeP. Moreover, what was more interesting is that this foursome Int-DeP's have identified gaps within the developing DT and PdM scientific literature. These were:

- Int-DeP 1: Follow the generic 4 phase plan for every asset, which guides, evaluates, and eliminates projects.
- Int-DeP 8: Implement first CBM, then transition to PdM strategy. Implement first CBM, then transition to PdM strategy.
- Int DeP 10: Capturing the tacit knowledge and preserving within the DT.
- Int-DeP 12: Identify generic configurable components for DT production that should become more affordable and faster deployed.

Int-DeP 1 proposed that no generic plan is provided on systematically transforming a set of assets to become more 'data-driven'. Int-DeP 8 contribution to the literature is to propose transitioning an organization from the CBM to a PdM strategy. Int-DeP 10 proposes that tacit knowledge capture in the DT is still an underdeveloped field of knowledge. Lastly, Int-DeP 12 proposes that no generic configurable DT components are identified to produce faster DT's.

The generic themes of the gap that identify DeP's are that these are more strategic, long-term, and on an organizational level. This explains the current status of the literature as these gaps identifying DeP's will be more of interest during the maturation phase of the literature while currently, it is in its infancy phase tackling other more relevant industrial and academic bottlenecks.

5.3. Data governance DeP

From section 4.5, the interviews derived the problems that RWS encountered with data governance. However, it was also concluded that RWS utilized no data governance best practices. Therefore, the literature found solutions as mentioned in section 2.2.4 to form Data Governance DeP's (DG-DeP's) that RWS can utilize for implementing DT and PdM for their AM at SSB.

DG-DeP 1:

The literature promotes the use of technologies to optimize the data life cycle, like HADOOP. The reasoning is that this type of software intelligently reduces and optimizes the stored big data that is ever-growing at RWS. This is a short-term solution for the still undetermined data life cycle. Currently it is saved for an infinite time until a decision (destroying the data) is made.

The DG-DeP: Utilize technologies to optimize the data life cycle management.

DG-DeP 2:

Sharing data with third parties (data farmers) solving problems is beneficial as RWS does not have the data analysis capabilities yet. However, this led to less control over the data. Using a data paragraph is an approach y RWS in the contracts. This is in line with the approach from literature as it suggests that to refocus from data ownership to on the data farmers can collect and manage the data in a more transparent and trusted approach.

The DG-DeP: Include a data paragraph in the contract which encompasses how the data farmer collects manages data in a transparent and trusted approach.

DG -DeP 3:

The SSB data is sensitive to terrorisms and thus is generally not shared with third parties. However, some form of data sharing should be allowed to use domain knowledge that RWs does not possess. A viable way is using a data-sharing system that identifies the leaks is using a knowledge base algorithm. This proactive security facilitates accountability of misusing parties like, for instance, an unauthorized user opening a classified map.

The DG-DeP: use a data-sharing system with reliable leaker identification to minimize the risks of sensitive SSB data sharing.

DG-DeP 4:

For RWS, it is unclear which roles are responsible for integrating data-driven technologies (data scientist, engineers etc.). From literature, data stewards, which are functions between IT and Business and are entrusted with the management of data, meta-data management and data governance are prescribed as roles. In other words, these are functions responsible for data-driven projects. The four functions are Business Steward (technical managers, business strategist etc.)overseeing the business side of the DDP, the technical steward (DT architects, data engineers, Process engineers) Responsible for the engineering side. These two will work together under the chief steward's supervision, who understands both activities. The last steward is the executive sponsor advising the firm's top management for long-term directions for transforming RWS from a pure civil organization to a hybrid organization by combining current asset management with information technology.

The DG-DeP: Incorporate 4 SSB transcending functions, which are responsible for overseeing the data related projects. These are executive sponsor, chief steward, business steward and technical steward.

5.4. Conclusion development of finalizing DeP’s

The input for this chapter were nine analyzed PdM and DT implementation case studies from chapter 2: literature study. These nine analyzed case studies were translated into 9 Lit-DeP’s and then combined with the 12 Int-DeP from chapter 4, resulting in 12 Ref-DeP's. Concurrently, literary solutions were provided for the current DG problems faced by RWS to form 4 DG-DeP's. These finalized 12 ref-DeP’s and 4 DG-DeP’s satisfies the main RQ1: ‘ What are the design principles for asset management?’
The following chapter uses these finalizing design principles to answer the main RQ2: ‘How can the design principles be used for asset management?’

6. Results – Testing the principles in practice using DFD analysis

From the previous chapter, a set of 12 finalized Ref-DeP's and 4 DG-DeP's were developed. In this chapter, these DeP’s are tested in practice using the Data Flow Diagrams (DFD)-analysis to see if the asset management is improved. This is executed by contextualizing the DeP’s within the DT and PdM technologies DFD’s. These DFD display systematically the improvements of the AM using the DeP’s.

The contextualization of DeP’s in the DFD of the PdM and DT goes as follows. Firstly, a custom-made DFD model is produced for both technologies: DFD-DT and DFD-PdM. These DFD models provide an 'ideal bird's eye view' of how the DeP's are applied. After that, the applicable DG-DeP's are overlayed on these models to increase the data quality leading to better usability of the models. At last, for the finalized DFD’s the opportunities and bottlenecks that an AM will face are presented, and insight is provided into how the principles helped improve the AM practice.
With the conclusions of this procedure of testing the principles in practice, the main RQ2 is answered: 'How can the design principles be used for asset management?'

6.1.1. DFD-DT (without DeP’s)

The first step of the DFD-DT development is building it from the insights provided by the DT architect at RWS. At this stage, the DeP's were not required, which resulted in the DFD-DT level 0 scheme shown in figure 6.1. It displays that the DT is related to three external entities: the operator interacting with the DT, SSB sensors providing sensor data, and the SSB actuators testing the pumps and compressors in the virtual environment. Level 0 without using DeP's is further expanded at DFD-DT level 1 as shown in figure D.1 in Appendix D.1. Generally, at level 1, it displays how the DT visualization is currently realized. Here, generalized processes are used: ETL, Databases (DB) and aggregation for continuously providing asset data visualization.
Some observations from the DFD-DT level 0 and 1 are that the DT is used in a one-dimensional and very basic way, which is primarily visualization. This visualization meets the minimum requirements to be a DT for extending the remaining asset lifetime, as mentioned in section 2.1.3. Fortunately, the DFD-DT's has a lot of potential by incorporating the short-term DT relevant ref-DeP's so that the DT can better contribute to a data-driven AM. This will be detailed in the next section.

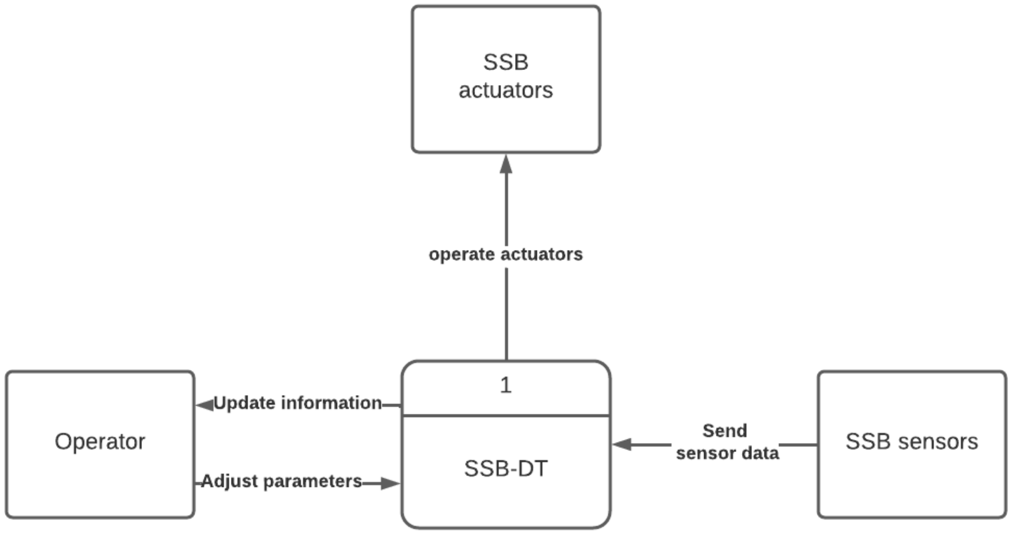


Figure 6. 1: Level 0 DFD-DT

6.2. DFD-DT

The first of the two models is the DFD-DT. This is developed by following three steps. Firstly, the DFD-DT is derived from the existing DT at SSB1. This is approached as this model will easily be transferable to other SSB's when built. Secondly, only the ref-DeP's that are short-term relevant and applicable for the DT from Appendix table C.3. are implemented. The reasoning for using short-term relevant ref-DeP's is because these are directly applicable and solve acute problems.
Lastly, the DG-DeP's which are applicable are introduced in the DFD-DT, resulting in the DFD-DT-DG, which improves the model usability.
After the fully developed DFD-DT and DFD-DT-DG models it is evaluated based on the expected future opportunities and bottlenecks.

6.2.1. DFD-DT (with Ref-DeP’s)

By incorporating the short-term, DT relevant ref-DeP's to the DFD-DT's from section 6.1.1. It further enhances it to better contribute towards a data-driven AM. These used DeP's are 3,4,6,7,9 and 11 as found in appendix table C.3. These are implemented to the DFD in the following manner:

- Ref-DeP 3: The currently used AM software is integrated into the DT. This is executed by programming directories on the virtual asset, which sent relevant sensor data to the AM software. The ERP, the second part of the ref-DeP, is not used as no branch-specific data is required, like production plans, orders, and availability of people for a DT.
- Ref-DeP 4: The model continuously compares changes between the simulated historic data and the real data received from the sensors and SCADA. The other part of the principle is not applicable as it is a procedure executed outside the DT model, like documenting the MoC.
- Ref-DeP 6: Data is provided by the 2 'external entity blocks' SCADA and SSB sensors.

- Ref-DeP 7: The actual DB is the primary storage of data. Nevertheless, when the data is older than one year, it will be transferred to the historic DB to minimize model lag time. The centralized and localized decision support is not relevant for the DT model.
- Ref-DeP 9: A knowledge DB stores prior problems solved happy, unhappy flows for operator training and problem-solution pair to propose solutions to the currently occurring situations.
- Ref-DeP 11: An Operator Training Simulation (OTS) is included, which uses the DT interface to learn users (external entity block), which are operators and maintenance workers 'to operate' the DT in a safe virtual environment.

Implementing the six ref-DeP as mentioned above resulted in the expansion of the level 1 DFD-DT with ref-DeP's 6 & 7 as shown in figure D.2 and adding a deeper layer, level 2, demonstrating ref-DeP's 3,4, 9 & 11 in figure D.3. With figure D.2 and D.3, sub-RQ2.1 is satisfied.

As a result of implementing these DeP's into the DFD-DT, the developed DFD-DT provides specific applications and insights which are found between the interacting DeP's. These are:

- **Efficient processing by data distribution based on lifetime:** Initially, all data raw data is stored in the actual DB. However, when the data is older than one year, it is aggregated by the ETL and then stored in the historic DB. This minimizes the model lag time as continuous visualization is updated. This contribution is found at level 1.
- **Additional data source:** Scada is added as it provides important computed data which the DT could use without using computing power. This contribution is found at level 1.
- **Knowledge base feeding the OTS:** Continuously updating the OTS with new happy and unhappy flow situations saved in the knowledge base. The input for the knowledge base could be artificially made situations or from a real closure of the SSB.
- **SSB sensor and SCADA feeding the AM software:** Sensor data is sent to the currently used AM software. These were previously coexisting but not linked, leading to latencies, inaccuracy, manual work errors and dependency on the worker's experience level. This implementation will significantly resolve the enumerated problems. This contribution is found at level 2.
- **Keeping DT up to date by automatically notifying the operator: The continuous validation will keep noticing the operator (update information),** suggesting that physical changes happened, so the DT requires synchronization with the updated physical assets. However, not every physical change will lead to detection by continuous validation. This feature serves as an addition to the manual MoC-activities. This contribution is found at level 2.

Concluding from the five specific applications, it is observed that these generally contribute to the DFD-DT by making it more multifunctional, especially by implementing an OTS and the linkage with AM software. This is in line with the critique for the current DFD-DT model, which seems to be one-dimensional and basic, as found in 6.1.1. In other words, the implementations of the short-term DT relevant DeP's that interact with each other provide a multifunctionally to the current way that the DT is used for AM.
The usability of this AM-multifunctional DFD-DT is improved by introducing the applicable DG-DeP's, which is discussed in the following section.

6.2.2. DFD-DT-DG

In this section, the DFD-DT from figures D.1 and D.2. are further enhanced by including data governance. This will improve the data quality leading to better usability of the data in the model. The used DG-DeP's are 1 and 3, as found in appendix C.4. The DG-DeP's are implemented in the following manner in the DFD-DT:

- DG-DeP 1: The open-source HADOOPsoftware is used as the intelligent software to optimizethe data life cycle of the historical and actual databases. Other intelligent softwares optimizing the data life cycle can also be used
- DG-DeP 3: An identification leaker is placed monitoring all the transferring sensitive data between the DT and the third parties. Misuse is notified, like copy-pasting classified files to a different server.

As a result of the DG-DeP's implementation, the DFD-DT-DG is produced and displayed in appendix D.1 figure D.4. By realizing the DG-DeP's, all the DB are optimized leading to less lag-time and facilitating on a controlled way of sharing data with external parties. For the latter, the introduction of external parties becomes more viable as they can improve the data quality using their resources and experiences. This controlled way of sharing is especially helpful for the SSB division as they are reserved and hesitant. This is because of concerns that terrorists may gain access to sensitive information. With the implementation of DG-DeP's, the usability of the model is positively affected. Thereby, sub-RQ2.2 is satisfied.

Opportunities and bottlenecks: By fully completing the DFD-DT-DG, the future opportunities and the bottlenecks that an AM will face are derived.

The major opportunity is to explore more existing AM activities that can be controlled and managed by the DT. This multifunctionality integration will make the DT the primary information exchange, decision-making support system, and strategizing tool. However, for such a promise, the DT needs to stay up to date for the whole virtual asset life cycle. This is achieved by making the last specific application mentioned in section 6.1.1 as a prerequisite when building and maintaining the DT's.
Another opportunity is to expand the proactive security of the DFD-DT-DG. An inflow of external knowledge becomes increasingly important because of the time pressure and increasing minimum requirements for AM. In this chapter, the identification leaker is presented to detect misuse of shared sensitive data but there are unexplored, better ways to secure the data sharing. Exploring this route extensively and setting up this system will boost the transition to a data-driven AM.

In addition, two social-related bottlenecks emerge. The first one is that the experienced decision-makers and operators mistrust the DT's supportive data-driven decision making. This mistrust can lead to minimal use of the DT and consequences to no improvement of AM while large investments are made. On the other side, beginning operators start to 'blindly accept' the DT's, slowing down the operator's asset knowledge development and relying too much on the 'multifunctional decision supporting' DT. This also leads to minimal improvement of the AM as operator knowledge is inevitable and is not replaceable by a decision-supporting model.
With this, the DeP's implementation and the associated insights for the DT is concluded. The following section will cover the implementation of DeP's for the DFD-PdM.

6.3. DFD-PdM

The second of the two models is the DFD-PdM. The DFD-PdM is developed in two steps. Firstly, the DFD-PdM is derived directly from the Ref-DeP's. This approach contrasts with DFD-DT, as RWS does not have an operating PdM from which the DFD can be inspired. Thereby the development of the DFD-PdM from ref-DeP's was necessary.
The second step is introducing the DG-DeP's which are applicable in the DFD-PdM, resulting in the DFD-PdM-DG, which improves the model usability.
After the fully developed DFD-PdM and DFD-PdM-DG models are evaluated based on the expected future opportunities and bottlenecks.

6.3.1. DFD-PdM (with Ref-DeP’s)

The first step of the DFD-PdM development is building it from the insights provided by the ref-DeP's. The used ref-DeP are 3, 4, 5, 6, 7 & 9 as found in appendix C.3. The ref-DeP's are used in the following manner to build the DFD-PdM:

- Ref-DeP 3: An adequate bidirectional information exchange is implemented. Instead of naming it Am-software, it is called ERP for this model as it also includes non-maintenance related specific data like production plans, orders, and availability of people.
- Ref-DeP 5: To schedule maintenance activities, the most important predictive maintenance parameters are included in the model, which is RUL repair and RUL replacement.
- Ref-DeP 6: data is provided by the 2 'external entity blocks', SCADA and SSB sensor.
- Ref-DeP 7: Separate decentralized and centralized decision support for maximum effective DM and minimizing lag time. While the local decision support is mostly automated, centralized decision support handles problems requiring a 'heavier' model and human interaction.
- Ref-DeP 9: use the knowledge DB To advise the local decision-support solutions for the associated problem.

In this respect, three levels are produced in the DFD-PdM using the five Ref-DeP's. At Level 0, it shows the overall relationship between the PdM model with the external entities: SCADA, sensors, operators and the asset management scheduling. This level is found in figure 6.6. Following, level 1 depicts how the main blocks in the PdM model contribute to data-driven asset management, as shown in appendix figure D.5. At this level, Ref-DeP's 3,6,7 and 9 are implemented. Lastly, the deepest level, level 2, expands the 'local decision-making support system' by including ref-DeP five as shown in appendix figure D.6. As a result of developing from the DeP's into the DFD-PdM, specific applications and insights are provided, which are found between the interacting DeP's. These are:

- **Three data source:** The DFD-PdM is described by starting from three different data sources: SCADA, sensors and historic DB. These will be further computed by the model. This contribution is found at level 1.
- **Implementation of Knowledge DB and support systems:** The incoming data is computed using the past experiences from the knowledge DB producing an RUL or transferring rare unsolvable situations to the centralized decision support system to produce an RUL.

- **RUL enhancement:** Thereafter, this RUL will be computed with the non-maintenance related information from the ERP, like availability of maintenance workers, funding and so on forth to build the 'replace RUL' and 'repair RUL'.
- **Meta-data production:** These types of RUL's will then be transferred to the asset management scheduling to which the operator can interact with making decisions. Data back these decisions to perform appropriate maintenance activities.

Concluding from the four specific applications, it is observed that these generally contribute to the DFD-PdM by taking many important factors to produce reliable automatic predictions based on data. This is especially the cause when the knowledge base and the centralized and decentralized decision supports interact to produce the best possible RUL. In other words, the implementations of the short-term DT relevant DeP's that interact with each other takes most of the important factors into account to predict better maintenance activities.
The usability of this DFD-PdM is improved by introducing the applicable DG-DeP's, which is discussed in the following section.

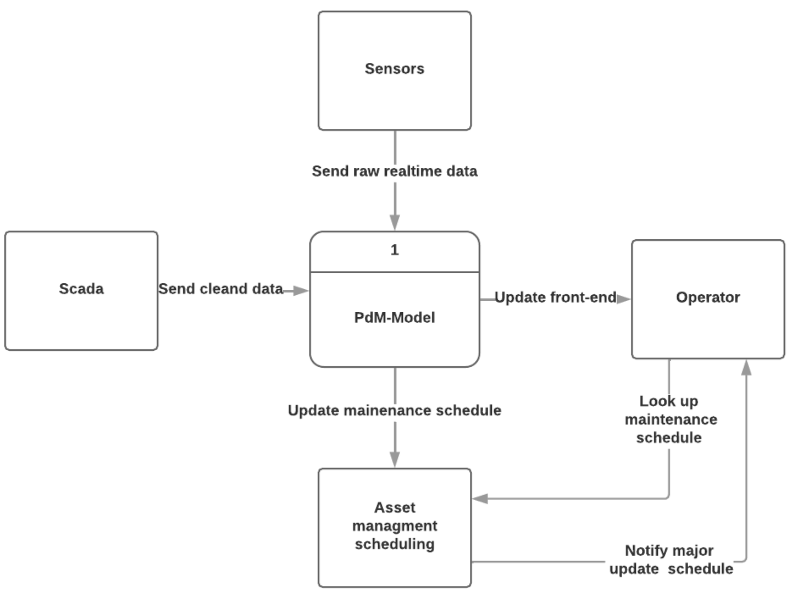


Figure 6. 2: DFD-PdM level 0

6.3.2. DFD-PdM-DG

The DFD-PdM in appendix figure D.5 is further enhanced by including data governance. The utilized DG-DeP's are 1,3 and 4, as found in appendix C.4. The DG-DeP's are implemented in the following manner in the DFD-PdM.

- DG-DeP 1: The open-source HADOOPsoftware is used as the intelligent software to optimizethe data life cycle of the historical and actual databases. Other intelligent softwares optimizing the data life cycle can also be used

- DG-DeP 3: An identification leaker monitors all the transferring sensitive data between the PdM and the third parties. Misuse is notified, like copy-pasting classified files to a different server.
- DG-DeP 4: The unsolved problems directed to the centralized decision support are solved by the data steward team, specifically the technical steward. In other words, the first line to solve data-driven problems will be the RWS internal data steward team. Moreover, they will be the contact person with the third party to evaluate the results offered by sharing the sensitive data.

As a result of the DG-DeP's implementation, the DFD-PdM-DG is produced as shown in appendix D.2 figure D.6. By realizing the DG-DeP's, all the DB are optimized, leading to less lag-time, facilitating a controlled way of sharing data with external parties and a dedicated team to maintain the models and solve difficult problems which reside between asset management and data. With the implementation of DG-DeP's, the usability of the model is positively affected. Thereby, sub-RQ2.2 is satisfied.

Opportunities and bottlenecks: The first opportunity is the expansion of the proactive security, which is similar to the opportunity for DT's mentioned and elaborated in section 6.1.2. The second opportunity is that implementing the proposed PdM model or something similar minimizes the human efforts. This positively affects to a better human resource management leading to many effective maintenance activities instead of periodic ineffective decision-making. The important process blocks are the localized decision support and knowledge DB.
As for the bottleneck, no severe constrictions are found except that building the knowledge DB will require time before the PdM model can operate automatically. With these identified opportunities and bottlenecks of the PdM and the DT (from section 6.1.3.), sub-RQ2.3 is satisfied.

6.3. Conclusions DFD-analysis

This chapter started with the inputs from the previous chapter: short-term relevant Ref-DeP's and the DG-DeP's. These DeP’s were tested in practice by contextualized within the DFD-DT and DFD-PdM to see if the asset management is improved. This is achieved by producing two custom made models relating the short-term relevant Ref-DeP's and the applicable DG-DeP:

- **DFD-DT-DG:**
    - The DFD-DT-DG has three layers utilizing the ref-DeP's 3,4,6,7,9 and 11.These ref-DeP’s enhanced the practice of using the DT from a primarily visualization into a more multifunctional usable digital asset for various asset management purposes.
      two of the AM purposes are the use of the OTS which is feed by the knowledge base and feeding the existing AM software with SSB sensor and SCADA data through the DT.
    - The DG-DeP’s increased the usability of the model by optimizing the data life cycle of the historical and actual databases (DG-DeP 1) and introducing proactive security, making sensitive data sharing with data capable oriented third parties possible (DG-DeP 3).
    - The opportunity from this model is to explore more existing AM activities that can be controlled and managed by the DT and expand the proactive security of the DFD-DT-DG, leading to an inflow of competencies from outside the organizational boundaries. The expected social-related bottlenecks are that experienced decision-makers and operators mistrust the data-driven support, leading to no use of the financial investment. On the

other hand, new beginning operators rely too much on the model, negatively affecting the AM. Importantly, operator knowledge is inevitable and is not replaceable by a decision-supporting model.

- **DFD-PdM-DG:**
  - The DFD-PdM-DG uses ref-DeP's 3,5,6 & 7 across three layers. These ref-DeP’s e enhanced the practice of the PdM by including the most important factors to produce reliable automatic predictions to better manage the maintenance scheduling. The important factors were the inclusion of Three data source, the implementation of Knowledge DB and support systems, the enhancement of the RUL with non-maintenance related information from the ERP and the production of suitable meta-data.
  - The usability of the model is increased by optimizing the data life cycle at the databases(DG-DeP 1), introducing proactive security (DG-DeP 3) and prescribing responsibilities for maintaining the models to the data stewards' teams (DG-DeP 4).
  - The opportunities identified are the expansion of the proactive security, similar to the opportunity for DT's and that implementing the proposed PdM model or something similar minimizes the human efforts. The latter leads to positively affecting to a better human resource management leading to many effective maintenance activities instead of periodic, ineffective decision-making.

By concluding this 'results chapter', all the insights and information are gathered to answer every main RQ's and sub-RQ's in the next chapter, conclusions.

7. Conclusions

This chapter provides the conclusion to this research project. The results are examined to answer the research questions in section 7.1. Consequently, the remarks on the findings are provided in section 7.2 and the limitations of the research are found in section 7.3. Recommendations for future research in section 7.4. Finally, the thesis concluded with and a brief reflection on the MoT Program in section 7.5.

7.1. Answering the research questions

The problem statement of this thesis was that there was no guidance for developing data-driven asset management while it urges to adapt because of operational and economical interferences.
Therefore, this research's goal (and contribution) was to develop design principles (DeP), which overcome real-world asset data-related problems during the implementation of DT, PdM and DG within the asset management context. From this problem statement followed up the research questions, which will be answered based on the examination of the results from this thesis.

RQ1: What are the design principles for asset management?

The design principles for asset management comprise two types of DeP’s. The first type is the Refined DeP’s (ref-DeP) prescribing on PdM and DT implementation within the AM-context. A set of 12 Ref-DeP's were developed by combining the Interview and literature-derived DeP's. The Ref-DeP's were further subdivided into short-term and long-term relevance. The short-term relevant Ref-DeP were easily applicable and further utilized in the DFD analysis. The long-term Ref-DeP's were more strategic in nature, thereby not further examined. The 12 Ref-DeP’s can be found in table 7.1.
The second type is the data governance DeP’s (DG-DeP) prescribing on data governance implementation within the AM-context. 4 DG-DeP’s were formed by combining the DG-problems from the semi-structured interview results and the literature derived solutions. The 4 DG-DeP’s can be found in table 7.2. These two types of DeP’s together assist in transitioning to a data-driven AM.

Ref-DeP #	Context of refined design principle	Long term/ short term
Ref-DeP 1	Follow the generic 4 phase plan for every asset, which guides, evaluates, and eliminates projects.	Long term
Ref-DeP 2	Organize periodic CoP to keep track of DT and PdM developments across divisions/SSB borders and optimize the knowledge management further by incorporating remaining knowledge processes: knowledge generation(1), knowledge integration (2) and collective project learning (3).	Long term
Ref-DeP 3	Use the DT as a user-friendly interface by making it compatible with the currently used AM software through directories, so human errors are reduced and integrate adequate information exchange between the shop floor and AM software to enhance the decision reactions.	Short term

Ref-DeP 4	Test the change first virtually in the DT and update, then make the physical change and finally, complete it with documenting in the MoC procedure, and concurrently validate the DT by continuous monitoring by comparing simulated data with the real plant to confirm that the real plant and DT is still identical.	Short term
Ref-DeP 5	Visualize the sensor data in temperature, pressure, RUL repair, RUL replacement, and other computed AM parameters to provide insight to facilitate maintenance planning.	Short term
Ref-DeP 6	Collect raw data at the source and the preprocessed data from SCADA and reuse it.	Short term
Ref-DeP 7	Store data based on the data life cycle and computed data based on complexity. data > 1 year goes to historical data, and easily computable data are solved by the decentralized level and harder computable data at the centralized level. This ensures the maximum effective DM and minimises lag time	Short term
Ref-DeP 8	Implement first CBM, then transition to PdM strategy.	Long term
Ref-DeP 9	Store all happy, unhappy flow and problem-solution pairs at the knowledge DB for reuse to prevent reinventing the wheel of solutions.	Short term
Ref-DeP 10	Capture tacit knowledge and preserving within the DT.	Long term
Ref-DeP 11	Facilitate operators and maintenance to learn 'to operate' in the existing DT environments in various SSB behaviours, which are: Normal and abnormal situations, Safety procedures and assembling and disassembling assets. This minimizes the operations and maintenance reaction time. .	Short-term
Ref-DeP 12	Identify generic configurable components to make DT production more affordable and faster deployed.	Long term

Table 7. 1: The final data refined design principles convers DT and PdM.

DG-DeP #	**Context of Data governance design principle**
DG-DeP 1	*Utilize technologies to optimize the data life cycle management.*
DG-DeP 2	*Include a data paragraph in the contract which encompasses how the data farmer collects manages data in a transparent and trusted approach.*
DG-DeP 3	*Use a data-sharing system with reliable leaker identification to minimize the risks of sensitive SSB data sharing.*

DG-DeP 4	*Incorporate 4 SSB transcending functions which are responsible for overseeing the data related projects. These are executive sponsor, chief steward, business steward and technical steward.*

Table 7. 2: The final data governance design principles.

1.1. What are the encountered problems hampering the DT and PdM implementation?

From the 33 effective semi-structured interviews, 12 encountered problems were derived. Every problem was unique and therefore listed outside the conclusions in appendix C.1.

What does SSB asset management demand from DT, PdM and data governance?

Three from the sex SSB were sampled. SSB1 was interested in DT-projects, especially how the DT could be more multifunctional, like implementing an operator training simulator.
SSB2 interest and priority lay in PdM technologies as some actions are already executed at SSB2 by hiring university interns for these projects.
SSB3 was not clear about its demand from PdM and DT. Instead, they stated that they would want a total solution provided by an external division within RWS.
The demand for data governance was not the priority for all the SSB's. The focus lay primarily on implementing DT and PdM technologies.

1.2. What are the best practices for the encountered problems and SSB demand to implement for DT and PdM?

For every PdM and DT encountered problems, a corresponding best practice was facilitated from the 33 effective semi-structured interviews. The 12 best practices are unique and therefore tabulated outside the conclusion in the Appendix in table C.1.

1.3. **What are the current data governance practices for DT and PdM implementation?**

Four data governance problems were derived from the semi-structured interviews. These were about the data life cycle, data ownership, sensitive data sharing and data stewardship. However, no solution or best practice was provided by the RWS-interviewees. Therefore, solutions from the data governance literature were proposed to RWS.

RQ2: How can the design principles be used for asset management?

The short-term refined design principles and the applicable data governance design principles were tested in practice within the DFD-DT-DG and DFD-PdM-DG.
The results of the DFD-DT-DG using the DeP's was that the model can be used multifunctionally for various AM purposes. The benefits of using DT in a multifunctional way is to reduce the fail change estimation. This is reached by implementing the OTS and increase decision support performance by connecting sensoring information to the currently used AM-software .
Moreover, the implementation of DeP's in the DFD-PdM-DG led to a model that includes the important factors to produce reliable automatic predictions. The benefits are that RWS has now a qualitatively approach model that could be used to develop their first PdM. The other benefits are that this qualitatively approach model will advise to perform maintenance only when needed.

What do the DFD-DT and DFD-PdM look like, based on the design principles?

The DFD-DT-DG has three layers (levels 0,1, and 2) utilizing the ref-DeP's 3,4,6,7,9 and 11. At level 1, the insights of Ref-DeP's 6,7 & 11 are found, while the remaining, Ref-DeP's 3,4 & 9, are found in level 2.

The DFD-PdM also has three layers: levels 0,1 & 2. Most short-term Ref-DeP (DeP's 3, 6, 7 & 9) are found at level 1. The only DeP, ref-DeP 5, can be found at level 2.

2.1. **How can data governance design principles be implemented within the DFD-DT and DFD-PdM to make the usability of PdM and DT more effective for AM purposes?**

Two DG-DeP's were implemented for the DFD-DT resulting in DFD-DT-DG. The DG integration increases the usability of the model by optimizing the data life cycle of the historical and actual databases (DG-DeP 1) and introducing proactive security, making sensitive data sharing with data capable oriented third parties possible (DG-DeP 3).

At the DFD-PdM, 3 DG-DeP's were used to produce DFD-PdM-DG. The usability of the model is increased by optimizing the data life cycle at the databases (DG-DeP 1), introducing proactive security (DG-DeP 3) and prescribing responsibilities for maintaining the models to the data stewards' teams (DG-DeP 4).

2.2. **What are the opportunities and bottlenecks found from the constructed DFD?**

The opportunities from the DFD-DT-DG are to explore more existing AM activities that can be controlled and managed by the DT (1) and expand the proactive security of the DFD-DT-DG (2), leading to an inflow of competencies from outside the organizational boundaries. The expected social-related bottlenecks are that experienced decision-makers and operators mistrust the data-driven support, leading to no use of this financial investment. On the other hand, beginning operators rely too much on the model, negatively affecting the AM. Importantly, operator knowledge is inevitable and is not replaceable by a decision-supporting model.

The opportunities identified from the DFD-PdM-DG are the expansion of the proactive security, similar to the opportunity for DT's and that implementing the proposed PdM model or something similar minimizes the human efforts. The latter leads to positively affecting to a better human resource management resulting in many effective maintenance activities instead of periodic, ineffective decision-making. Further, no severe constrictions are found except that building the knowledge DB will require time before the PdM model can operate automatically.

7.2. Remark on the findings

The following highlights unexpected or exciting results that arose during the investigation to the research questions:

- **Academic relevant DeP's:** By developing Ref-DeP's, it was noticed that four Int-DeP's had no associated Lit-DeP. Therefore, these Ref-DeP's were similarly formulated as the corresponding Int-DeP. Moreover, more interesting is that this four some Int-DeP's have identified gaps within the developing DT and PdM scientific literature.
  Int-DeP 1 proposed that no generic plan is provided for systematically transforming a set of assets to become more 'data-driven'. Int-DeP 8 contribution to the literature is to propose how

to transition an organization from the CBM to a PdM strategy. Int-DeP 10 proposes that tacit knowledge capture in the DT is still an underdeveloped field of knowledge. Lastly, Int-DeP 12 proposes that no generic configurable DT components are identified to produce faster DT's.
The generic themes of the gap that identify DeP's are that these are more strategic, long-term, and on an organizational level. This explains the current status of the literature as these gaps identifying DeP's will be more of interest during the maturation phase of the literature while currently, it is in its infancy phase tackling other more relevant industrial and academic bottlenecks.

- **Lack of RWS DG-DePs:** RWS 4 DG-problems but unexpected did not have data governance best practices. The semi-structured interviews gave a 'potential demand' for the 4 DG- problem, which was not further used in this thesis.
- **Long-term ref-DeP's:** During further subdividing the Ref-DeP's, a set of Long-term Ref-DeP was strategic from nature and less applicable. Therefore, these DeP's were not further utilized within the DFD analysis.
- **SSB3 demand:** Unexpectedly, SSB3 could not pinpoint DT or PdM demand. Instead, they wanted a complete solution provided by another RWS-division before SS3 would consider investing its resources in these data-driven technologies. This resulted in a sample that did not contribute to this thesis.
- **Contrasting views**: The contrasting views from the semi-structured interviews were the SSB openness for new technology, the alignment between AM priority and the data-driven technology, and data access by third parties. These contrasting viewpoints unknowingly construct invisible impermeable walls which prevent knowledge spillovers, resulting in departments maturing at different rates and worsening the understanding and communication between them at RWS.

7.3. Limitation of the research

Some limitations and assumptions are made to deliver an insightful report within the given time constraint.

- **DFD verification:** The DFD's introduced in chapter 6 are verified but not implemented nor put into practice in the RWS asset management. Therefore, a validation step of the models is lacking.

- **Limitation of literature refinement:** The literature approach within this thesis assumes that the assets are continuously used like in a chemical plant. Consequently, all papers are assumed to be for '24/7 continuous operated' assets. However, the SSB assets are parsley operated, averaging 1 to 2 closings a year. This mismatch between continuous operation and parsley operation leads to different asset degradation curves and thus limits the insights of theoretical background.

- **Confined sampling:** The semi-structured interviews had 2 major confinements that limited this research sampling. The first is that the brainstorm sessions were only held at three of the six SSB. Which means the remaining three SSB's demand for the data-driven technologies were not takes into account in this thesis. The second is that the semi-structured interviews were held only at the SSB and the lock complex devisions, despite the fact that RWS has more division (tunnels, bridges and so on forth) which are gradually implementing data-driven techologies. Therefore, the insight of the other RWS-devisions were excluded from this thesis.

- **Further analysis of remaining DeP's:** Not every (Ref- or DG-)DeP were further analyzed for opportunities and bottlenecks. The two unexamined groups were 'Long-term Ref-DeP's' and DG-DeP 2 & 4.

7.4. Recommendations for future research

To further develop the body of knowledge of 'data-driven asset management' using the insights of this thesis, the managerial recommendations are twofold:

- **Invest earlier in data governance practices**: By utilizing early on data governance during the PdM and DT implementation, intersecting problems like data life cycle, data ownership, sensitive data sharing and data stewardship are resolved before the data governance issue becomes unmanageable when every SSB has it on DT and PdM model. This early implementation will reduce future ad-hoc problem solving that will inevitably originate from the novel digital assets.

- **Utilize solutions for 'SSB transcending knowledge transfer':** As many DDP are concurrently developing, a system should be used to centralize the knowledge across all SSB's. Here, the long-term Ref-DeP's play their part by suggesting an integrated strategic approach. For instance, organize periodic CoP's to keep track of DT and PdM developments across divisions/SSB (ref-DeP 2) and establish a plan to strategically guide, evaluates, and eliminate projects. (ref-DeP 1) and so on forth.

A lack of these two will inhibit the RWS's long-term vision of transitioning from a pure civil organization to a hybrid organization: A synergy between civil engineering and data-driven technologies.

7.5. Reflection on the learning process

Reflecting on the learnings from this thesis, I was in a unique position in learning a qualitative approach alternatively of the standard quantitative approach, which every engineer and researcher with a technical background uses for solving problems. I learned that the qualitative approach has pros in providing early insights on socio-technological problem sets.

While learning this new approach, I would have made two decisions differently. Firstly, I would have emphasized the academic contribution from the beginning phase. Due to the nature of this research, a focus was put on a practical approach, thereby unjustly neglecting the work for the academic contribution. This, unfortunately, led to the elongation of the thesis.
Secondly, a better arrangement had to be made regarding the deliverables. For this research, two conceptually the same reports had to be written: A thesis report in English for the academic public and an advice report in Dutch for RWS had. Writing these two similar reports took a significant portion of my time in the writing process. This resulted in a lower reporting quality and less attention paid to the reasoning of the arguments. Fortunately, In the last month, it was decided that the thesis report would be sufficient in addition with a presentation at RWS, and no advice report was anymore required.

In continuation, a reflection is also made on the contribution of MoT program to this thesis. From the MoT program primarily two courses were used: Research methods and Business process management & Technology. The research method provided clear guidance on building the research methodology for which an entire chapter was dedicated in this thesis. The business process management & Technology helped to think systematically of useful schemes during the DFD design.
However, a recommendation for the MoT-program is to employ more compulsory courses related to ‘data as an asset as it is an increasingly important domain in the coming decades. An example is I&C Architecture Design which probably would have helped me in this thesis.

Bibliography

Adam Barone, D. K. (2022, January). *Definition assets*. Investopedia.

Alladi, B. S., & Prasad, S. (2018). Big data life cycle: Security issues, challenges, threat and security model. *International Journal of Engineering and Technology(UAE)*, *7*(1.3 Special Issue 3), 100–103. https://doi.org/10.14419/ijet.v7i1.3.9666

Bamberg Andreas, U. L. B. S. B. M. K. N. (2021). *The Digital Twin-Your Ingenious Companion for Process Engineering and Smart Production*. https://doi.org/10.1002/ceat.202000562

Brous, P., Janssen, M., & Herder, P. (2019). Internet of Things adoption for reconfiguring decision-making processes in asset management. *Business Process Management Journal*, *25*(3), 495–511. https://doi.org/10.1108/BPMJ-11-2017-0328

Chen, J., Chee, ·, Lim, P., Kim, ·, Tan, H., Govindan, · Kannan, & Kumar, A. (2021). *Artificial intelligence-based human-centric decision support framework: an application to predictive maintenance in asset management under pandemic environments*. https://doi.org/10.1007/s10479-021-04373-w

Cheng, J. C. P., Chen, W., Chen, K., & Wang, Q. (2020). Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Automation in Construction*, *112*. https://doi.org/10.1016/j.autcon.2020.103087

Choubey, S., Benton, R., & Johnsten, T. (2019). *Prescriptive Equipment Maintenance: A Framework; Prescriptive Equipment Maintenance: A Framework*.

Ellixson, A., & Griffin, T. (2016). *Farm Data: Ownership and Protections*. https://ssrn.com/abstract=2839811Electroniccopyavailableat:https://ssrn.com/abstract=2839811

Faiz, R. B., & Edirisinghe, E. A. (2009). Decision Making for Predictive Maintenance in Asset Information Management. In *Interdisciplinary Journal of Information, Knowledge, and Management* (Vol. 4).

Fong, P. S. W. (2003). Knowledge creation in multidisciplinary project teams: An empirical study of the processes and their dynamic interrelationships. *International Journal of Project Management*, *21*(7), 479–486. https://doi.org/10.1016/S0263-7863(03)00047-4

Groba, C., Cech, S., Rosenthal, F., & Gössling, A. (2007). *Architecture of a Predictive Maintenance Framework*.

Havard, V., Jeanne, B., Lacomblez, M., & Baudry, D. (2019). *Production & Manufacturing Research Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations*. https://doi.org/10.1080/21693277.2019.1660283

Herrington, A., Herrington, J., Mantei, J., Herrington, A., & Herrington, J. (2009). *Design principles for mobile learning Design principles for mobile learning Recommended Citation Recommended Citation Design principles for mobile learning*.

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. In *Scandinavian Journal of Information Systems* (Vol. 19, Issue 2).

Kim, D., Lee, S., Kim, D., & Cano, J. C. (2021). *An Applicable Predictive Maintenance Framework for the Absence of Run-to-Failure Data*. https://doi.org/10.3390/app11115180

Kuan Cheong, L., Chang, V., & Kuan, L. (2007). *Association for Information Systems AIS Electronic Library (AISeL) The Need for Data Governance: A Case Study Recommended Citation "The Need for Data Governance: A Case Study" The Need for Data Governance: A Case Study*. http://aisel.aisnet.org/acis2007/100

Leingang, C., Oppelt, M., Pfeiffer, B.-M., Institute of Electrical and Electronics Engineers, IEEE Industrial Electronics Society, & Universidad de Zaragoza. Aragón Institute for Engineering Research. (2019). *Proceedings, 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) : Paraninfo Building, University of Zaragoza, Zaragoza, Spain, 10-13 September, 2019*.

Overheid.nl. (2022, January). *Water law*.

Panfilov, P., & Katona, A. (2018). Building Predictive Maintenance Framework for Smart Environment Application Systems. *9TH DAAAMINTERNATIONAL SYMPOSIUM ON INTELLIGENT MANUFACTURING AND AUTOMATION*. https://doi.org/10.2507/29th.daaam.proceedings.xxx

Rahul, K., & Banyal, R. K. (2020). Data Life Cycle Management in Big Data Analytics. *Procedia Computer Science*, *173*, 364–371. https://doi.org/10.1016/j.procs.2020.06.042

Schranz, C., Strohmeier, F., & Damjanovic-Behrendt, V. (2020). A Digital Twin Prototype for Product Lifecycle Data Management. *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, *2020-November*. https://doi.org/10.1109/AICCSA50499.2020.9316506

Thomas Vollmer & Robert Schmitt. (2015). *Integrated shop floor data management for increasing energy and resource efficiency in manufacturing*.

Uma Sekaran, R. B. (2018). *Research methods for business* (7th ed.). John Wiley & Sons.

Vonk, B., Klerk, W. J., Fröhle, P., Gersonius, B., Heijer, F. den, Jordan, P., Ciocan, U. R., Rijke, J., Sayers, P., & Ashley, R. (2020). Adaptive asset management for flood protection: The FAIR framework in action. *Infrastructures*, *5*(12), 1–16. https://doi.org/10.3390/infrastructures5120109

Wang, F., & Hannafin, M. J. (2005). *Design-Based Research and Technology-Enhanced Learning Environments*.

Wiseman Associate Professor of Law, L., Sanderson, J., & Robb, L. (2018). Rethinking Ag Data Ownership. In *Farm Policy Journal |* (Vol. 15, Issue 1). https://www.commerce.senate.gov/public/index.

Zhang, Z., Guan, Y., Ma, X., & Zhang, L. (2019). *AuditShare: Sensitive Data Sharing with Reliable Leaker Identification*. http://arxiv.org/abs/1907.11833

Zipper, H., Auris, F., Strahilov, A., & Paul, M. (2018). Keeping the digital twin up-to-date - Process monitoring to identify changes in a plant. *Proceedings of the IEEE International Conference on Industrial Technology*, *2018-February*, 1592–1597. https://doi.org/10.1109/ICIT.2018.8352419

Appendix

A. Operationalization

To answer the research questions, the abstract concepts are defined in the literature in section 2.1. should be reduced to measurable and tangible variables to perform the semi-structured interviews. The operationalization from concepts to interview questions are elaborated in this section A.1. afterwards, the construction of the questionnaire in three main sections is found in A.2.

A.1. Operationalization of abstract concepts

From section 2.1. four abstract concepts are described from the literature. These will be operationalized to make measurable, tangible interview questions.

Asset management:
The main factors defining asset management are found in section 2.1.1. To operationalize this concept, questions can be asked about the overall performance, risks and investment cost and how these factors are balanced currently. This is an excellent introductory question, which helps to understand the underlying organizational theme in RWS before narrowing it down to data governance questions.

IQ 10: How is the overall performance, risks and investment costs balanced within asset management/maintenance?

Predictive maintenance:
Outlining the maintenance strategies is synthesized for this research to pinpoint the progression of various initiatives and determine the extent of PdM maturation perceived by the RWS personnel. This will incentivize urgency to fasten the developments or confirm that their current implementation plan is on track.
The validity is safeguarded by including the entire domain around maintenance strategies found in the literature.

Digital twin:
For operationalizing the DT, two concepts were mentioned in the literature study. The first is the integration stage of DT. The integration stage classifies various DT's and implies how far the development of DT is. Questions that could be asked are the characteristics of the DT model they use. The second concept is the relationship between DT and PM. Here, the contribution of DT to PM can be asked to measure the variations in knowledge by comparing across various SSB.

Q1: What type of DT or PM projects are you associated with?

Q2: What is the impact of this data-driven project on the current RWS working methods at the management and shopfloor levels?

Q3: How does this project benefit maintenance activities? How is the efficiency of the maintenance measured?

Q4: Was it required to associate commercial companies with this project?

Q5: To what extent is the data insight for asset management? How is the efficiency of maintenance measured?

Q6: How can DT or PM contribute/fulfil the need for SSB asset management in the future?

Q7: Is the SSB open for collaborations with commercial companies?

Q8: How do you view training and education in the future that will include PM & DT?

Q9: Do you experience difficulty communicating about data-driven technologies outside your team?

Data governance:
For digitization, the focus is put on the technical side. However, utilizing emerging technologies within an incumbent firm requires looking at the governance. Here, various questions can be asked about key terms in data governance, which is new and unknown. Questions will be asked about important terms, like who is responsible for these developments or the developments around data access, lifecycle, and data quality?

Q11: How are the current policies and procedures defined for proactive and effective data management?

Q12: Who is the initiator/responsible/team for digitization of asset management at your department

Q13: What is the general role of data as an asset within RWS?

Q14: How is the data managed in terms of data quality, - Access & - life cycle?

A.2. Constructing the interview questions from operationalized concepts

The questionnaire is built up of 3 main sections: The first section questions regard DT and PM for the asset management and their developments in their field. The second part covers data governance and its developments. The last section rounds off the Interview by closing questions like asking recommended interviewees, additional documentation, and questions regarding this research.

Section 1 of Interview: Data-driven technologies questions

The first four questions are dedicated to focusing group 1. The questions explore relevant initiatives from the lock complexes that may interest SSB asset management.

> Q1: What type of DT or PM projects are you associated with?
>
> Q2: What is the impact of this data-driven project on the current RWS working methods at the management and shopfloor levels?
>
> Q3: How does this project benefit maintenance activities? How is the efficiency of the maintenance measured?
>
> Q4: Was it required to associate commercial companies with this project?

The following three questions are asked to focus group 2. PM and DT concepts per SSB are mapped out to understand their current stage of development and their future perspectives per SSB. The answers per SSB will be compared with each other.

> Q5: To what extent is the data insight for asset management? How is the efficiency of maintenance measured?
>
> Q6: How can DT or PM contribute/fulfil the need for SSB asset management in the future?
>
> Q7: Is the SSB open for collaborations with commercial companies?

The following two questions target both focus groups. Ideas around learning and developments are shared.

> Q8: How do you view training and education in the future that will include PM & DT?
>
> Q9: Do you experience difficulty communicating about data-driven technologies outside your team?

Section 2 of Interview: Data governance questions

Question-related to both focus groups. Current data governance practices are sought.

> Q10: How is the overall performance, risks and investment costs balanced within asset management/maintenance?
>
> Q11: How are the current policies and procedures defined for proactive and effective data management?
>
> Q12: Who is the initiator/responsible/team for digitization of asset management at your department
>
> Q13: What is the general role of data as an asset within RWS?
>
> Q14: How is the data managed in terms of data quality, - Access & - life cycle?

Section 3 of Interview: concluding Interview

These are closing questions to round off the Interview and seek new potent interviewees.

> Q15: Do you know other essential stakeholders I can interview regarding data-driven tech: PM, DT?
>
> Q16: Are there related documentation that could provide more information?
>
> Q17: Do you know other essential stakeholders I can interview regarding data governance?
>
> Q18: Do you have some questions?

B. Interviewees

interviewee #	Function	Related DPP	SSB demand: Brainstorm session	Department
1	Lead Architect A-omgeving MK	1, 4, 5, 6		CIV
2	Senior asset advisor	-		-
3	Senior asset advisor	6		CIV
5	Pilots: Smart sensoring + Digital twinning	2		WNN
6	Maintenaince engineer: vital asset	3		PPO /Z&D
7	Manager program Vital assets	1,2		CIV
8	Senior Asset Management/product owner	-		MN
9	PdM bij het programma R-BIM	-		CIV
10	Advisor SSB industrial automation	4,6		CIV
11	Program Manager: vital assets	1		CIV
12	Advisor Smart Maintenance ON	3		MN,
13	Technical manager	-		CIV
14	Technical advisor	3		PPO
15	Technical advisor	3		PPO
16	Application manager	-		CIV
17	Rijks trainee	7		CIV
18	Manager SSB3	7		ZD
19	Manager SSB3	7	X	ZD
20	Engineer SSB3	7	X	PPO
21	Engineer SSB3	7	X	CIV
22	Senior-advisor asset management	1		ZD
23	SSB 3 advisor	7		WNZ
24	SSB 3 advisor	7		WNZ
25	Solution architect maritime management	3,5		
26	SSB1 maintenance engineer	5	SSB1	MN
27	Advisor	-		NOVA
28	Advisor	-		NOVA
29	Advisor	-		NOVA
30	Enterprise architect expert	7	SSB3	CIV
31	Fail chance manager SSB 1	5	SSB1	MN
32	BiK program manager	4,5,6,7		CIV
33	SSB manager (SSB 1)	5		MN

Table B. 1: The list of effective interviews with the RWS personnel

C. Design principles

C.1. Interview derived DeP (Int-DeP)

Int-DeP #	Problem	Best practice	SSB Demand	Int-DeP	DDP #	Relevance (DT/PdM/ Both)	Long term/ short term
1	No step-by-step actionable plan that is generic and applicable for every asset to make it data-driven within RWS.	A generic 4 phase plan with in-between go/no-go that is applicable for every asset.	-	Follow for every asset the generic 4 phase plan, which guides, evaluates and eliminates projects.	1	Both	Long term
2	Difficult to track multiple projects concurrently and communicate effectively to evaluate redundancy of projects and share in-field practices.	Recurring organized CoP to share, analyze, and transform learning experiences in implementing data-driven technologies into tools, specific instructions, training and preventing redundant projects.	-	Organize periodic CoP to keep track of DT and PdM developments across divisions/SSB borders.	1	Both	Long term
3	The human error of Incorrect asset fault registration influences the maintenance schedule negatively.	Use the DT as the user-friendly interface for translating physical asset information to the existing AM software by clickable directories per	SSB1: making the existing DT invest more diverse and multifunctional.	Use the DT as a user-friendly interface By making it compatible with the currently used AM software through directories so	2	DT	Short term

		asset(componen t).		human errors are reduced.			
4	Physical asset is ‘updated’ by replacing machines and minor changes, but the virtual assets become ‘outdated’ and not relevant for AM purposes.	Implementing DT updates during the MoC procedures.	SSB1 & SSB2: Test the change first in the DT in various scenarios before making the physical change.	Test the change first virtually in the DT and update, then make the physical change and finally, complete it with documenting in the MoC procedure.	2	DT	Short term
5	No insights on the historic and current status of the assets leading to ‘over maintenance’ and expensive reactive maintenance.	Visualize sensor data (pressure, temperature, and computed AM parameters) in PowerBI, which facilitates better AM decision-making.	-	Visualize sensor data (pressure, temperature, and computed AM parameters) in PowerBI, which facilitates better AM decision-making.	3	PdM	Short term
6	Unclear from which sources the raw data can be extracted.	Use raw sensor data and preprocessed data from SCADA.	-	Find raw data from the source and preprocessed data from the SCADA.	3	Both	Short term

7	Storing all the data but not much information is relevant leading to slowing down of the model.	Data older than one year is aggregated to the desired size and labelled as historic. Data < 1 year stays full and complete.	-	Create two separate databases (DB): Historical DB holding aggregated data >1 year and the primary database: holding untouched complete data < 1 year.	3	Both	Short term
8	Transitioning to PdM is a giant step with many complications and uncertainties.	Implement CBM before considering the PdM strategy.	-	Implement first CBM, then transition to PdM strategy.	3	PdM	Long term
9	Uncertainties arise regarding the correct interaction between the new OS and SSB, which leads to an undesired increased fail chance estimation.	Building a virtual environment (DT) where the OS can be validated in happy and unhappy flow with speed up functions before commissioning it to the real SSB	-	Use a database storing various happy and unhappy flow scenarios to test the new OS and cure the OS teething problems	4	Both	Short term
10	The tacit knowledge for understanding the logics disappears every 15 years leading to tedious redesigning the wheel (OS) over again.	Capturing the tacit knowledge and preserving within the DT.	-	Capture tacit knowledge and preserving within the DT.	4	DT	Long term

11	Operators have too little practical experience with closing the physical SSB and, therefore, undertrained.	Facilitate a simulator (virtual environment) where operators can be trained in a didactic manner in happy and unhappy flow scenarios.	Demand from SSB1, 2. Integrating also happy and unhappy flow for maintenance	Facilitate training operations and maintenance programs within the DT environment.	5	DT	Short-term
12	Designing the DT takes an enormous amount of time to commission fully.	Build/use generic configurable components based on the previous DT projects.	-	Identify generic configurable components to make DT production more affordable and faster deployed.	6	Both	Long term

Table C. 1: The systematic derivation of DeP from analyzing DDP.

C.2. Literature derived DeP (Lit-DeP)

Lit-DeP #	Problem	Resulting case study	Lit-DeP
1	managing multiple teams across the firm introduces redundant projects, concentrated knowledge within one division and minimal knowledge transparency.	Incorporate the boundary-crossing, knowledge generation and collective project learning when managing knowledge within a origination , division or even within a program	Take various knowledge-creating processes into account when managing knowledge which is: Boundary crossing(1), knowledge generation (2), knowledge integration (3), knowledge sharing (4) and collective project learning (5)
2	The lack of upper-level information within the virtual asset results in time-consuming manual synchronization.	Enhanced and fine-tuned the current virtual asset by feeding the virtual asset with the machine and upper enterprise-level information , which will facilitate enhanced appropriate decision reactions.	Integrate an adequate information exchange between the shop floor and ERP to facilitate enhanced appropriate decision reactions.

3	Virtual assets could not shorten and minimize the negative effects of turnarounds because it diverges from the physically changing plant. Caused by maintenance and process optimizations cause physical and logical changes.	Updating the model modification through continuous monitoring by comparing simulated data with the real plant to confirm that the real plant and DT are still identical.	Updating the virtual assets through continuous monitoring by comparing simulated data with the real plant to confirm that the real plant and DT are still identical.
4	it is unclear what and how the results from the PdM will visualize in a concrete manner to improve the AM decision-making.	Simplifies the PdM output for maintenance planning in two estimated RUL per asset: repair and replacement.	Visualize the results in the repair and replacement RUL's to facilitate maintenance planning
5	it is unclear what and how the results from the PdM will visualize in a concrete manner to improve the AM decision-making.	RUL is a standard that is generally applied in PdM literature.	Use the remaining Useful Lifetime (RUL) as additional AM indicators.
6	It is unclear which sources should be employed to feed the PdM model	used the PdM model for further predictive analytics, SCADA and sensors, and reusing it. The sensors generate unclean, incomplete, and low-quality data, which must be cleansed and processed. SCADA data is typically pre-filtered; thus, it requires minimal processing.	Collect raw data at the source and preprocessed data at SCADA and reuse it.
7	Regard a common practice is to send all data acquisition to a central data warehouse. The result of this strategy is centralized decision-making. However, this strategy goes against the idea of sending the right information at the right time from the correct location because all data is centralized while all the	As a result, a combination of localized and centralized decision support systems should be utilized. Here, the local decision support will operate as the main server.	Save all the data at decentralized DB and necessary data at Centralized DB for effective DM

	equipment are decentralized over an area. As a result, haltering the suitable choices and slowing down the decision speed. (Choubey et al., 2019) (Kim et al., 2021)		
8	Similar problems reoccur, requiring the same solution. However, the 'same' solution is reproduced in an ad-hoc and iterative manner. It is meaningless to come up with new solutions to a problem that has already been solved successfully. An efficient way should be utilized.	The attained knowledge from the PdM model and the pairing of problem to solutions (referred to as problem-solution pairs) can be stored in the knowledge database, so when the next time a problem is discovered or reported, this knowledge base can be utilized first to find a solution, before generating a possible new solution to this problem.	First, find solutions from the knowledge DB for known problem-solution pairs before generating handmade novel solutions.
9	A DT is not always evident to operators and contractors how it can offer value for business models. This is especially true for the contribution to operator training.	Use the DT to learn to 'operate' in normal and abnormal situations, refresh safety procedures, understand the assembling and dissembling of assets in order to minimize the reaction time with a safe environment where undesired consequences are acceptable, and the operations and maintenance are taught in a didactical manner.	Train operators in the OTS to understand operations and maintenance in a didactical manner to minimize reaction time by the following training sessions: - Normal and abnormal situations. - Safety procedures - Assembling and disassembling assets

Table C. 2: The synthesis of Lit-DeP by analyzing the problem and the associated best practice that are prescribed from the scientific literature.

C.3. Refined DeP (Ref-DeP)

Ref-DeP #	Interview derived Problem	Int-DeP	Lit-DeP	Ref-DeP	Applicability (DT/PdM or both)	Relevance (Long term/ short term)
1	No generic instructions to transition to make assets data-driven	Follow the generic 4 phase plan for every asset, which guides, evaluates, and eliminates projects.	None	Follow the generic 4 phase plan for every asset, which guides, evaluates, and eliminates projects.	Both	Long-term
2	Difficult to track multiple projects concurrently and communicate effectively to evaluate redundancy and share in-field practices.	Organize periodic CoP to keep track of DT and PdM developments across divisions/SSB borders.	Take various knowledge-creating processes into account when managing knowledge which is: Boundary crossing(1), knowledge generation (2), knowledge integration (3), knowledge sharing (4) and collective project learning (5).	Organize periodic CoP to keep track of DT and PdM developments across divisions/SSB borders and optimize the knowledge management further by incorporating remaining knowledge processes: knowledge generation(1), knowledge integration (2) and collective project learning (3).	Both	Long-term

3	The human error of Incorrect asset fault registration influences the maintenance schedule negatively.	Use the DT as a user-friendly interface By making it compatible with the currently used AM software through directories, so human errors are reduced.	Integrate an adequate information exchange between the shop floor and ERP to facilitate enhanced appropriate decision reactions.	Use the DT as a user-friendly interface by making it compatible with the currently used AM software through directories, so human errors are reduced and integrate adequate information exchange between the shop floor and AM software to enhance the decision reactions.	Both	Short-term
4	Physical assets are 'updated' by replacing machines and minor changes, but the virtual assets become 'outdated' and not relevant for AM purposes.	Test the change first virtually in the DT and update, then make the physical change and finally, complete it with documenting in the MoC.	Updating the virtual assets through continuous monitoring by comparing simulated data with the real plant to confirm that the real plant and DT are still identical.	Test the change first virtually in the DT and update, then make the physical change and finally, complete it with documenting in the MoC procedure, and concurrently validate the DT by continuous monitoring by comparing	DT	Short-term

				simulated data with the real plant to confirm that the real plant and DT is still identical.		
5	No insights on the historic and current status of the assets leading to 'over maintenance' and expensive reactive maintenance.	Visualize sensor data (pressure, temperature, and computed AM parameters) in PowerBI, which facilitates better AM decision-making.	Lit-DeP 4: Visualize the results in the repair and replacement RUL's to facilitate the maintenance planning Lit-DeP 5: R Use the Remaining Useful Lifetime (RUL) as additional AM indicators.	Visualize the sensor data in temperature, pressure, RUL repair, RUL replacement, and other computed AM parameters to provide insight to facilitate maintenance planning.	PdM	Short-term
6	Unclear from which sources the raw data can be extracted.	Find raw data from the source and preprocessed data from the SCADA.	Collect raw data at the source and preprocessed data at SCADA and reuse it.	Collect raw data at the source and the preprocessed data from SCADA and reuse it.	Both	Short-term
7	Storing every sensor data Will slow down the model.	Create two separate databases (DB): Historical DB holding aggregated data >1 year and the primary database: holding untouched complete data < 1 year.	Save all the data at decentralized DB and complex necessary data at centralized DB for effective DM.	Store data based on the data life cycle and computed data based on complexity. data > 1 year goes to historical data, and easily computable data are solved by the	Both	Short-term

				decentralized level and harder computable data at the centralized level. This ensures the maximum effective DM and minimises lag time		
8	Transitioning to PdM is a giant step with many complications and uncertainties.	Implement first CBM, then transition to PdM strategy.	None	Implement first CBM, then transition to PdM strategy.	PdM	Long-term
9	Uncertainties arise regarding the correct interaction between the new operating system(OS) and SSB, which leads to an undesired increased fail chance estimation.	Use a database storing various happy and unhappy flow scenarios to test the new OS and cure the OS teething problems	First, find solutions from the knowledge DB for known problem-solution pairs before generating handmade novel solutions.	Store all happy, unhappy flow and problem-solution pairs at the knowledge DB for reuse to prevent reinventing the wheel of solutions.	Both	Short-term
10	The tacit knowledge for understanding the logics disappears every 15 years leading to tedious redesigning the wheel (OS) over again.	Capturing the tacit knowledge and preserving within the DT.	None	Capture tacit knowledge and preserving within the DT.	DT	Long-term
11	Operators have too little practical experience with	Facilitate training programs for operators	Train operators in the OTS to understand operations and maintenance in a	Facilitate operators and maintenance to learn ‘to	DT	Short-term

	closing the physical SSB and, therefore, undertrained.	within the DT environment.	didactical manner to minimize reaction time by the following training sessions: - Normal and abnormal situations. - Safety procedures - Assembling and disassembling assets	operate' in the existing DT environments in various SSB behaviours, which are: Normal and abnormal situations, Safety procedures and assembling and disassembling assets. This minimizes the operations and maintenance reaction time. .		
12	Designing the DT takes an enormous amount of time to commission fully.	Build/use generic configurable components based on the previous DT projects.	None	Identify generic configurable components to make DT production more affordable and faster deployed.	Both	Long-term

Table C. 3: The synthesis of refined DeP by combining the interview and literature derived DeP's.

C.4. Data governance DeP (DG-DeP)

DG-DeP#	Interview derived problem	Interview derived potential demand	Resulting case study	DG-DeP
1	De data lifetime cycle is undetermined; therefore, all data is saved for an infinite time	An efficient way to store the data.	Promote using technologies for optimizing the data life cycle. The reasoning is that this software intelligently reduces and optimizes big data.	*Utilize technologies to optimize the data life cycle management.*

	period till a decision is made.			
2	All the ownership of the data is at a third-party company which lead to lower control on the asset as data.	Clarification of implementation of legislation paragraph regarding data ownership in the contract	Refocus from data ownership to how the data farmers can collect and manage the data in a more transparent and trusted approach.	*Include a data paragraph in the contract which encompasses how the data farmer collects manages data in a transparent and trusted approach.*
3	Due to the SSB information being sensitive information for terrorism, third party collaborations are censed	Some form of data access should be allowed as some domain knowledge is RWS does not possess.	A viable way is to use a data-sharing system that identifies the leaks using a knowledge base algorithm. This proactive security facilitates accountability of misusing parties like, for instance, an unauthorized user opening a classified map.	*Use a data-sharing system with reliable leaker identification to minimize the risks of sensitive SSB data sharing.*
4	Unclear roles for integrating data-driven technologies for the AM	Hiring data scientists, engineers etc.	From literature, data stewards are functions between IT and Business and are entrusted with the management of data, metadata management and data governance are prescribed as roles. In other words, these are functions that are responsible for data-driven projects. The four functions are executive sponsor, chief steward, business steward and technical steward.	*Incorporate 4 SSB transcending functions which are responsible for overseeing the data related projects. These are executive sponsor, chief steward, business steward and technical steward.*

Table C. 4: The synthesis of DG-DeP by solving interview derived problems by literature derived case studies.

D. DFD-models

D.1. Digital twin DFD

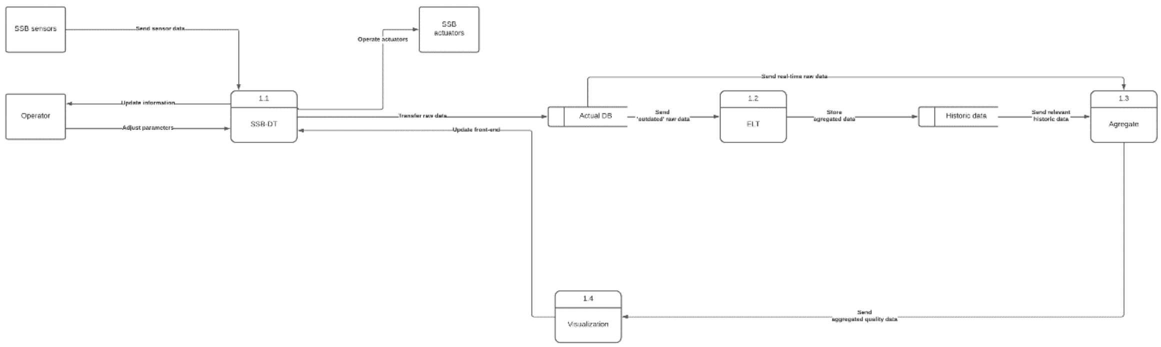


Figure D. 1: level 1 DFD-DT excluding the DeP's

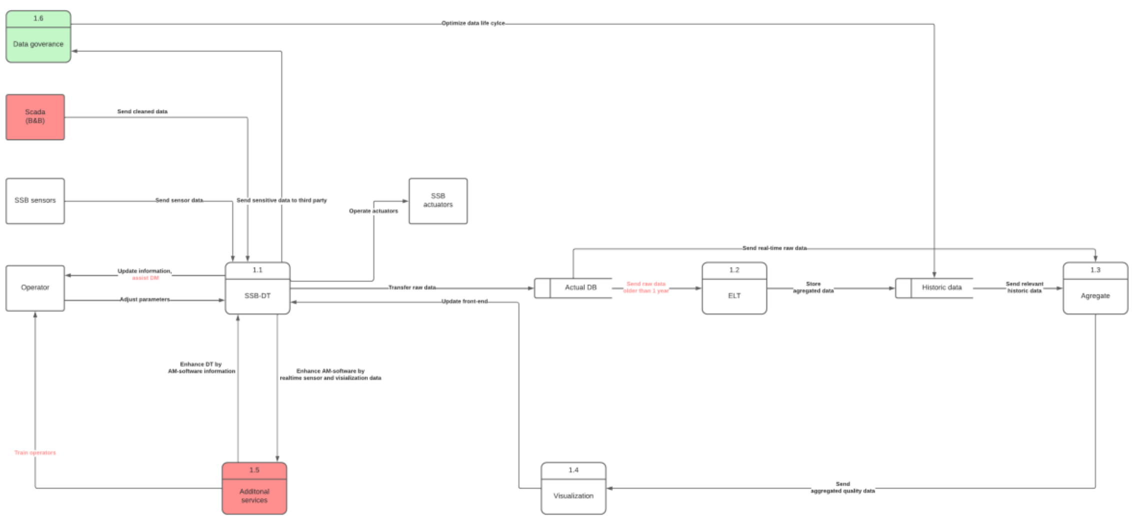


Figure D. 2: Level 1 DFD-DT including general DG-DeP's and Ref-DeP blocks

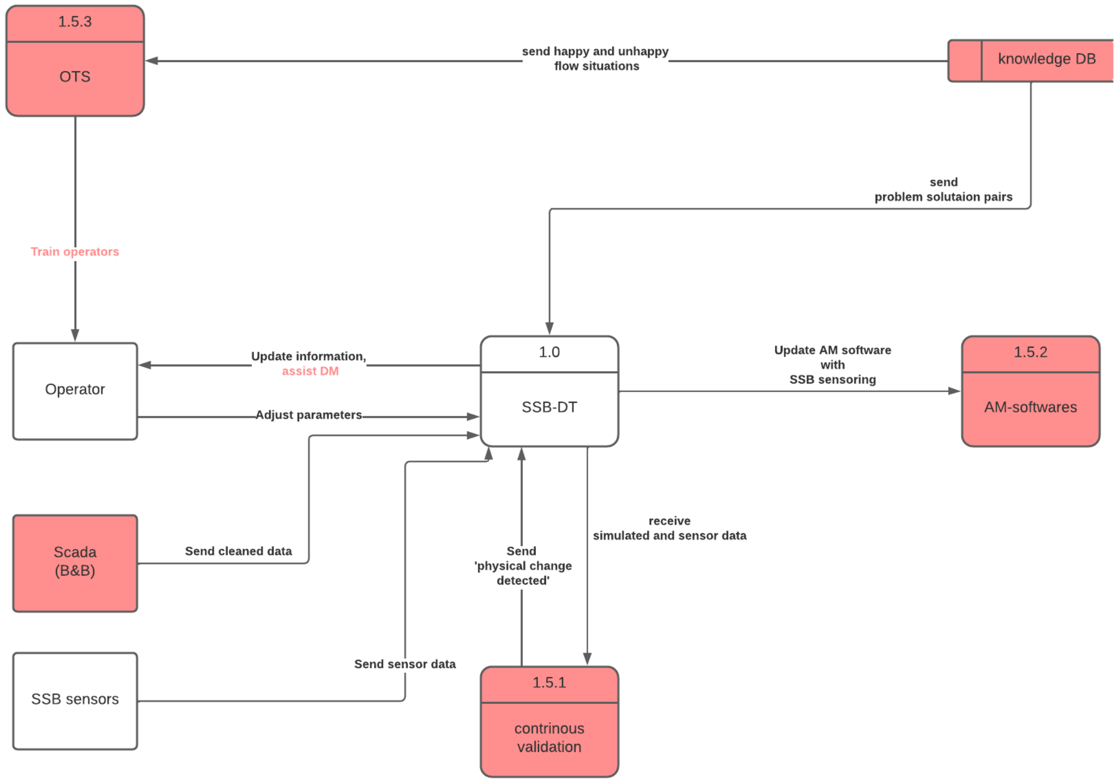
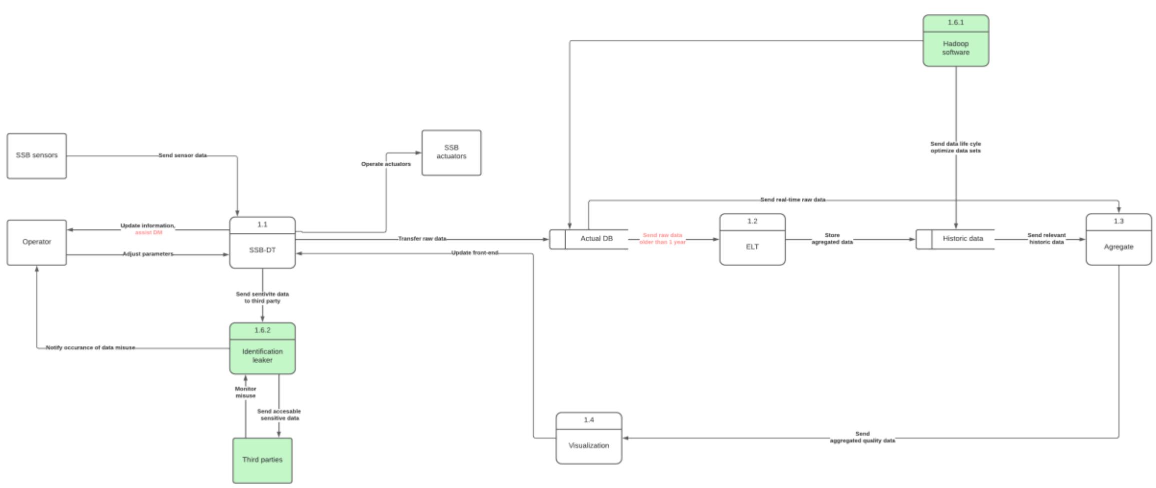


Figure D. 3: Level 2 DFD-DT focusing on Ref-DeP's



s

Figure D. 4: Level 2 DFD-DT focusing on DG-DeP's

D.2. Predictive maintenance DFD

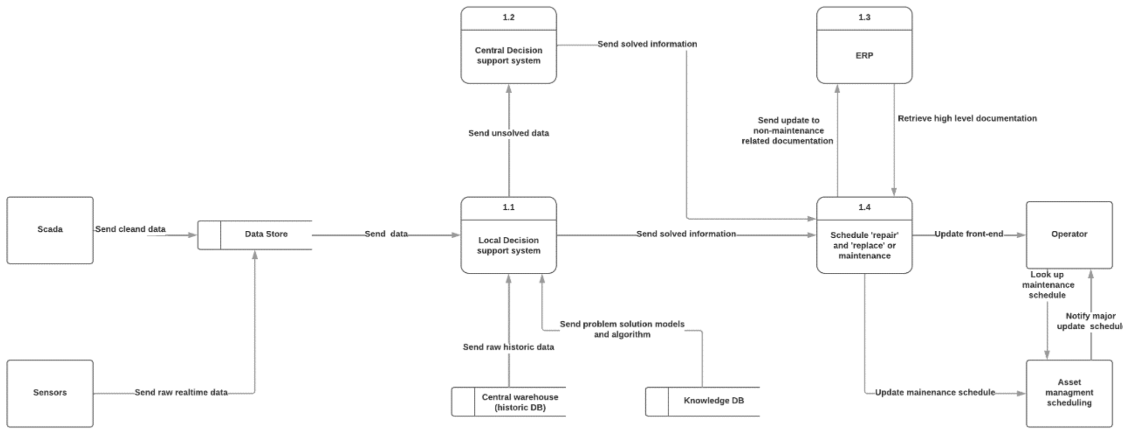


Figure D. 5: Level 1 DFD-PdM including Ref-DeP's

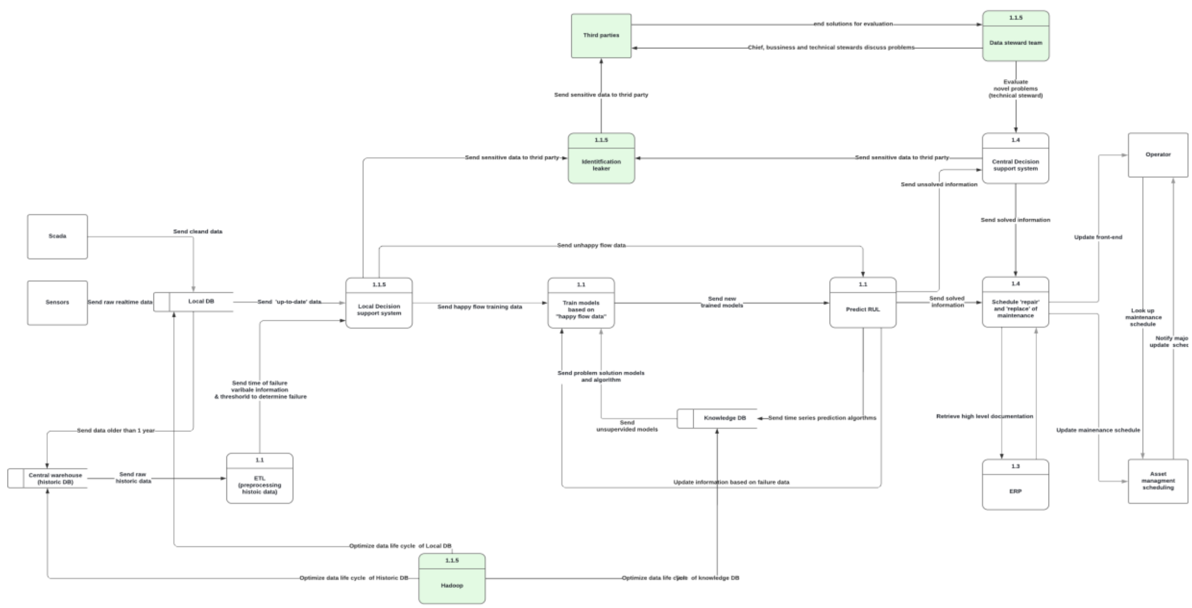


Figure D. 6: level 2 PdM including data governance (green). The process block 1.1.: local decision support from level 1 is expanded.