FACULTY OF SCIENCE AND TECHNOLOGY

MASTER THESIS

Study programme / specialisation:
Industrial Asset Management

The spring semester, 2022

Open

Author: Andreas Sandvik Jakobsen

Andreas Sandvik Jakobsen
(signature author)

Course coordinator: Jayantha Prasanna Liyanage

Internal Supervisor: Jayantha Prasanna Liyanage
External Supervisor: Ingrid Sternhoff (Aker BP)

Thesis title: Study of DataOps as a concept for Aker BP to enable data-driven assets.

Credits (ECTS): 30

Keywords:
DataOps
Data Management
Data Governance
Asset Management
Data Assets

Pages: 81

+ appendix: 0

Stavanger, 15.06.2022

Abstract

The oil and gas industry have faced many obstacles over the decades and in recent years the industry has had to endure increased pressure from the market, a global crisis caused by the coronavirus and oil prices reaching an all time low. Low oil prices stimulate a reaction from the industry having to do more with less, finding new ways of working and exploring previously untapped opportunities for improvement through digitalization.

Aker BP aim to be at the forefront of digitizing the exploration and production (E&P) industry and their digital transformation is more than just technology. Aker BP aim to build digital capabilities, develop digital mind sets and implement new ways of working where decisions are driven by data. The company show a willingness to experiment with and develop new technologies, however data management efforts are increasing as digitalization projects are realised from Aker BP's digital lab. An investment into data management is therefore required to allow for automation over time and ensure the right competence within Aker BP.

In collaboration with Aker BP, this thesis investigates the importance of an organization wide data management and data governance strategy aligned with business objectives and how the emerging concept of DataOps can enable Aker BP's ambitions of becoming a data-driven company. DataOps is an emerging approach advocated by data practitioners to cater to the challenges in data analytics projects. The thesis examine how Aker BP work with data in their organization today through interviews and discussions with different parts of the organization and assess how the principles and concepts of DataOps can be applied. To answer this, a DataOps maturity model has been developed and Aker BP's ways of working evaluated by using the model. The principles and concepts of DataOps have then been considered and an implementation plan and critical success factors for succeeding with DataOps or maturing data management and data governance efforts are presented.

The thesis focuses on the principle that DataOps' goal is to liberate data from its sources to its consumers and proposes the necessary steps to embark on the DataOps journey and maturing data management within the company. The research show that data management efforts are limited and based on ad-hoc needs and fast changing priorities. The pressure on the data management function to verify, clean, liberate, analyse and advise on data is growing and the participants in the research emphasize the need for change. Based on the DataOps maturity model developed, the steps required to increase Aker BP's DataOps maturity are identified. Further research should include implementation of the identified steps and investigate the efforts to further automating business processes and maturing Aker BP's ability to work with data.

Acknowledgments

I would like to express my gratitude to my faculty supervisor Jayantha Prasanna Liyanage for his supervision, guidance and sharing of knowledge and insight during this project. I would also like to thank Ingrid Sternhoff, external supervisor at Aker BP, for the continuous support. Your guidance and feedback throughout this process, in addition to your expertise and experience from industrial projects have helped me see challenges from a different perspective.

I wish to give special recognition to the Aker BP employees and colleagues who I have been in contact with for your support and engagement. The support, technical knowledge and experiences from the industry provided have been critical to the findings in this thesis. I would also like to thank my friends and family who supported me and offered their insights into the study.

Without your time and assistance, completion of this thesis would have been unachievable.

Contents

List of Figures

List of Abbreviations

AI Artificial Intelligence.

API Application Programming Interface.

AR Augmented Reality.

BMS Business Management System.

BU Business Unit.

CCTV Closed-Circuit Television.

CD Continuous Delivery.

CDF Cognite Data Fusion.

CI Continuous Integration.

CMMI Capability Maturity Model Integration.

D&W Drilling & Wells.

DAMA Data Management Association.

DataOps Data Operations.

DBT Digital Business Transformation.

DGO Data Governance Office.

Dr.Q Data Reliability & Quality.

E&P exploration and production.

ExpRes Exploration & Reservoir Development.

GIS Geographic Information System.

GUI Graphical User Interface.

I&P Improvement & Performance.

IAM Intelligent Asset Management.

IIoT Industrial Internet of Things.

IM&T Information Management & Technology.

IoT Internet of Things.

ISO International Organization for Standardization.

KPI Key Performance Indicators.

MDM Master Data Management.

ML Machine Learning.

NCS Norwegian Continental Shelf.

NORSOK the Norwegian shelf's competitive position.

NUI Normally Unattended Installation.

O&G oil and gas.

OAD Operations & Asset Development.

OSDU Open Subsurface Data Universe.

PDS Petrotechnical Data Systems.

PO Product Owner.

PRO Projects.

SaaS Software as a Service.

SME Subject Matter Expert.

VR Virtual Reality.

WITSML Well-Site Information Transfer Standard Markup Language.

Chapter 1

Introduction & Overview

This study aims to aims to analyze how Aker BP will build the leading independent offshore exploration and production (E&P) company in Norway through data driven digital transformation and a DataOps approach to data analytics and operations. The functional attributes and main principles of DataOps will be described and connected to the operational layers of the Drilling & Wells department within Aker BP. This introductory section provides the background for the study, an examination of why Aker BP emphasize digitalization and data in their strategy and a presentation of the research questions which the thesis intend to answer.

1.1 Aker BP ASA

In 2016 BP Norge AS and Det norske oljeselskap ASA announced the creation of Aker BP ASA, an independent oil and gas company combining the assets and expertise from both companies' Norwegian E&P operations with an ambition to become the leading independent offshore E&P company. At the end of 2021 a merger between Lundin Energy and Aker BP was announced where Aker BP will acquire Lundin Energy's oil and gas related activities. The company will become the largest listed E&P company focused purely on the Norwegian Continental Shelf (NCS) and the combined company will be characterized by increased scale, world-class quality and high returns with a workforce of approximately 2300 employees

[2].

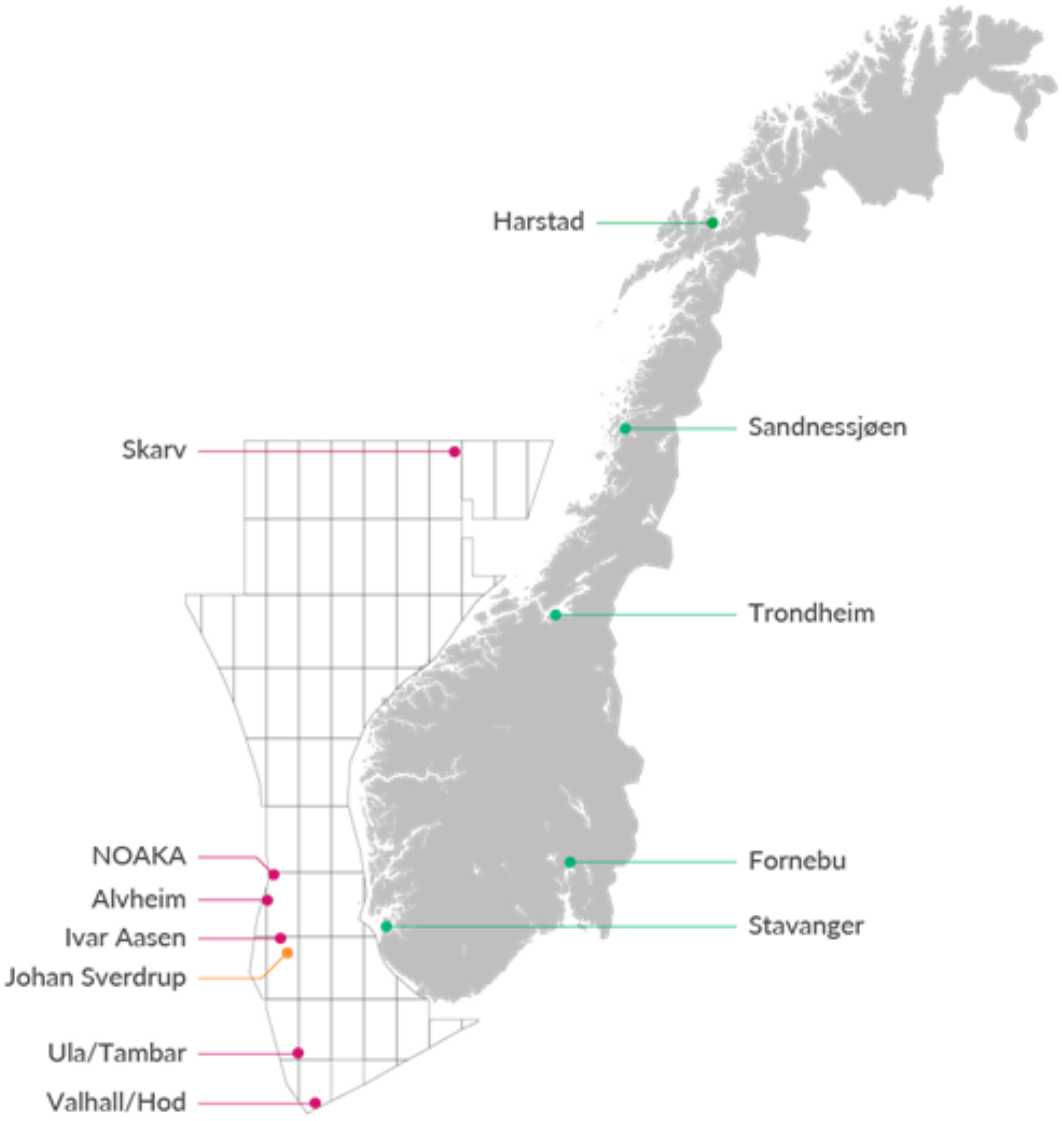


Figure 1.1: Aker BP's Assets on the Norwegian Continental Shelf

1.2 Background

As stated, the vision of Aker BP ASA is to become the leading independent offshore Exploration & Production (E&P) company. The company looks to do so by exploring the previously untapped opportunities for improvement through digitalization by being in the forefront of digitizing the E&P industry. To succeed with this one of the strategic priorities is to become a data-driven company by changing the way they work with industrial data, reducing costs from exploration to development and optimize production with algorithms, automation and digital tools.

As digitalization is a key to improvement and has been on Aker BP's agenda since its origin. The digital program EurekaX is Aker BP's hub for transformation and business optimisation, where cross-functional teams work to solve specific tasks. Through EurekaX, Aker BP has taken a leading role in the digitalization of the oil and gas industry which had led to technology evolving at unprecedented rates. However, with increased technological capabilities it has become imperative to the industry to put data at the center. To become a data driven company and harvest the value of information, it is key that data is seen as an asset on its own and managed accordingly. Important aspects to consider will be to ensure data is upheld to certain standards in the organization that makes it shareable and reusable, that data can be easily accessed from outside of the domain where it was located and domain data teams consider their data assets as products and consumers of the data as customers.

Access to data, across departments and functions, enables the ability to make better decisions. However, it has been revealed that Aker BP's operationalization of roles for data handling have not been good enough. The company believe there is great potential for improvement in maintaining and systematically developing digital competence in Aker BP. In the current state, data silos are reducing the availability and accessibility of data and trust in data is limited due to missing consistency in data management and governance. This is affirmed by a lack of enterprise-wide standards and guidelines on data that are defined, implemented, and actively used in the organization, resulting in inconsistency and inefficiency. Aker BP wants to mature its data management and governance to maximize utilization of data, get more insight

out of data, and reduce inefficiencies caused by lack of trust in data [4].

With increasing environmental and sustainability pressures, the oil and gas (O&G) industry must do more with less. Aker BP's success relies heavily on its approach to sustainability to meet the environmental corporate social responsibility standards set by society. Aker BP aim to create substantial long-term value from the natural resources we have available, thus enabling society to restructure the economy and introduce measures that can halt climate change. In addition, the industry will contribute knowledge, data and experience to new industries, creating growth and improvement beyond the oil and gas business [3]. Aker BP has a clear strategy of low carbon low cost and seek opportunities to optimise efficiency in their production. The realisation of these optimisation opportunities will result in lowered emission and cost savings. A majority of the optimisation opportunities in Aker BP focus on liberating sensor data from their operations and further contextualizing the data to enable data driven decisions. Further increasing the importance of an organization-wide data and information management, or DataOps, strategy aligned with business objectives.

1.3 Scope & Objectives

This thesis explores one of the key frameworks Aker BP aim to deploy when working with data and analytics to become a data-driven company, namely DataOps. When placing data at the center and attempting to maximize business value from analytical insight Aker BP has realised that the organization needs to mature its data capabilities, culture and processes. The main problem to investigate in the thesis is how DataOps as a framework and concept can support Aker BP's strategic priority to become a data-driven company. It will do so by analyzing Aker BP's maturity when working with data and compare it to a maturity model developed as part of the research on DataOps. Based on the maturity stage Aker BP qualifies for, an implementation plan and critical success factors for increasing Aker BP's DataOps maturity to the next stage is presented. The thesis outcome will provide additional support and authority on DataOps approaches and practices that will be implemented in Aker BP to reduce the cost and time to handle complex data.

The objectives of the thesis are to:

1. Provide an introduction to DataOps and develop a generic DataOps maturity model.
2. Review the processes and areas described in the maturity model in Aker BP today.
3. Create a brief overview of potential data management or DataOps tools that will be relevant for Aker BP.
4. Discuss the concepts of DataOps that will be important to Aker BP to increase their DataOps maturity.
5. Prepare an analysis of what steps need to be taken to implement a DataOps approach and the critical success factors related to the implementation.

1.4 Methodology

To achieve the objectives listed in section 1.3 and explore the efforts required to utilize DataOps as a concept for Aker BP to become more data-driven the thesis paper follows an exploratory and qualitative research approach. A literature study of data management practices and concepts, in addition to a comprehensive study of DataOps articles from DataOps enthusiasts such as Datakitchen, Cognite and IBM reflect the qualitative aspect of the research process. The exploratory research approach investigates the research questions in a qualitative manner and has been used to study data management and DataOps practices within Aker BP where there are few preexisting studies conducted on the topic. Collecting information on the topic has been performed by formulating the hypothesis that DataOps will be a key enabler in supporting Aker BP's strategic priority to become a data-driven company and investigating current practices to understand how the DataOps concepts can be implemented in practice going forward. The research has been performed in close collaboration with the Aker BP organization and the researcher has been present at the Aker BP offices during the writing of this thesis to investigate the practices in detail. Meetings and discussions with Aker BP stakeholders have been conducted throughout the period of writing and the specific DataOps approach for Aker BP

has been developed in iterations based on internal changes and continuous feedback from the organization. The choice of working in iterations in close collaboration with the organization was chosen as DataOps is a term and concept foreign to most, and initial meetings showed a large enthusiasm for the concept albeit with little contribution on how it can be beneficial for Aker BP. As the concept has matured over this period in Aker BP, the feedback on the topic of the thesis has also become more prolific. To be able to compare the results from the analysis of the current practice and the concept to be developed, a DataOps maturity model was developed through experience and observation with the goal of being able to link the two phases. The groups involved in discussions on the topic were mainly from the four groups of data management, digital transformation (EurekaX), the business (D&W operations) and the newly established data governance office in Aker BP. The sample size of people involved from each group varied from 2-6 participants and discussions were both performed in informal meetings, and in structured working sessions. The opening set of questions to initiate the research were in the form of:

- What do you think is required for Aker BP to become more data-driven?
- What are the key areas to succeed on this?
- What are the tools and processes that can help Aker BP get there?
- What role does your group have in a company driven by analytical decisions and data-driven insight?
- How can Aker BP utilize a DataOps function to reach their strategic priorities?

1.5 Delimitation

The research conducted will be delimited to the operations business unit of D&W in Aker BP and the implementation plan will be related to this business unit. Therefore, whilst the concepts and suggestions proposed are relevant to the other operational business units of Operations & Asset Development, Exploration & Reserovoirs and Projects in Aker BP they will not be part of the research interviews. The focus will be on DataOps principles and how Aker BP can mature their ability to work with data at the maturity they currently. Creating data pipelines and automating

data monitoring is not covered in detail. The thesis focuses on the necessary steps needed to mature from stage 2 to 3 in the DataOps maturity model and the steps needed to go from stage 3 to 5 are not covered in detail this thesis.

Chapter 2

Theoretical Review

2.1 Industrial Digitalization

Digitalization, as described by [60], is "a process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies". In order to exploit the potential of Industry 4.0 and Intelligent Asset Management (IAM), industrial organizations have to introduce changes in their technology infrastructures and their methodologies, but also in their managerial decision making. [40]. Digital transformation of industrial processes is currently driving the next generation of asset management solutions that are destined to be intelligent, automated and cost-efficient. Nevertheless, the transition from current paradigms to Industry 4.0 compliant, fully digital approaches to asset management is an extremely challenging task. It entails changes in the technological infrastructure, the asset management processes and the business management of industrial organizations. Therefore, understanding the challenges and identifying possible solutions across all these directions is the key to a successful transition industrial digitalization [27]

Industrial Digitalization commonly refer to Industry 4.0 technologies. Notably, 3D printing, advanced robotics, Internet of Things (IoT), Augmented Reality (AR), Virtual Reality (VR) and digital twins are becoming increasingly important to most industries. Global oil and gas companies are seeking to optimize their digital trans-

formation for efficient and safe operations and to achieve sustainability goals by leveraging Artificial Intelligence (AI) and IoT solutions [54]. By leveraging AI, for instance, oil and gas companies can predict and address leakages in near-real-time as well as reduce the number of wells drilled while maintaining high production or output levels. At the same time, IoT sensors and robotics can be leveraged to improve worker safety and the technologies are already driving digital trends across the oil and gas sector [54].

The emergence of Industry 4.0 enables the collection and processing of large amounts of data for individual assets and their lifecycle processes. For example, based on predictive analytics, Industry 4.0 can derive insights on how to optimize the maintenance of an asset [27]. Likewise, using sensor data it's possible to create accurate life cycle assessments for an asset, perform simulations and optimize processes like well design and drilling. However, leveraging data-driven insight also imposes methodological changes in the collection and processing of datasets. Many miss the opportunity to extract the true value of their data because information is housed and maintained in different silos across the organization. These silos are often managed independently by different user groups, leading to data inconsistency, inefficiency, and cost [11]. Consolidating, understanding, and analyzing these disparate datasets may not be possible or take excessive time and capacity, and errors in the data are found too late in the lifecycle leading to both inaccuracies and inefficiencies in decision making.

2.1.1 Digital Twins

Gartner defines a Digital Twin as "a digital representation of a real-world entity or system. The implementation of a digital twin is an encapsulated software object or model that mirrors a unique physical object, process, organization, person or other abstraction. Data from multiple digital twins can be aggregated for a composite view across a number of real-world entities, such as a power plant or a city, and their related processes" [24]. They capture important information about an asset such as engineering content, maintenance history, operating parameters and sensor data. Aside from the obvious benefit of creating a single, secure repository for all asset documentation, digital twins can positively impact the operational efficiency,

reliability and agility of oil and gas manufacturers [22]. By adopting digital twins, organizations will be able to gather real-time data feeds from sensors to evaluate an asset's state and condition or simulate the drilling of a well for training by applying solutions powered by AI, ML and modelling. The increased adoption of digital twins and smart data systems is described to be key to oil and gas operator's success [28]. To achieve this centralizing data management and further having proactive, robust data quality controls will ensure downstream systems, such as a digital twin, are provided with clean, accurate information in a timely manner. The importance of investing in managing data is quickly becoming evident and automating these efforts can free up resources to focus less on data issues and more on data analysis and insights that form meaningful analytical decision [11]

2.2 Engineers Assets & Unattended Facilities

Since production started in 1971, oil and gas have been produced from a total of 119 fields on the Norwegian shelf. At the end of 2021, 94 fields were in production: 71 in the North Sea, 21 in the Norwegian Sea and two in the Barents Sea [48]. The offshore oil and gas industry have contributed significantly to economic growth in Norway, both for the Norwegian welfare state and through repercussions to other industries. The upstream, also referred to as exploration and production (E&P), segment in the oil and gas industry involve complex assets such as drilling rigs, floating production units and production platforms. These assets consist of a variety of essential systems for safety and security such as fire and gas detection systems, intercom and alarm systems, Closed-Circuit Television (CCTV), satellite communication systems and multiple control systems enabling safe and efficient exploration and production processes. The industry is embracing the Industry 4.0 era and its adaption to oil and gas has brought advanced digitalization of assets with implementation of industrial digital twins, use of cloud technology, Industrial Internet of Things (IIoT), Augmented Reality (AR), Virtual Reality (VR), artificial intelligence, remote testing and control and automated systems [54].

As the technologies above are developing at an unprecedented rate, the O&G industry experiencing an increase of remote work, predictive maintenance and unattended

platforms. Normally Unattended Installation (NUI), also refered to as Normally Unmanned Installations, are offshore facilities primarly designed to be operated remotely through automated processes and without the presence of personnel [20]. As more inspection data is collected from assets the industry will have an unique opportunity to integrate immersive virtual reality technology and 3D simulations into surveyor or operational training sessions [53]. Traditional asset monitoring relies on gathering data from third-party sources and onboard equipment, operators are then tasked with combining data sources from various sensors or assets and define the actions to be taken. These processes are time-consuming and the adoption of digital twin technology is addressing this problem by offering a deeper understanding of assets with interactive and real-time views of the asset's condition [53]. Developing assets where daily operations are controlled by onshore teams have clear benefits of reduced costs, improved safety and increased efficiency [53]. However, these digital advancements and remote operations require a solid foundation of data and robustness in data gathering solutions.

2.3 Data Governance

Data Management Association (DAMA) [34] define Data Governance as 'The exercise of authority, control, and shared decision making (planning, monitoring and enforcement) over the management of data assets'. Data Governance defines who within an organization has the authority and control over data assets and how those data assets can be used. It can also be described as a function that supports an organization's overarching data management strategy and a framework to provide a holistic approach to collecting, securing, managing, and storing data [7]. The full potential of data management can only be realized after an organization has established clear data governance, and can lead to improved decision making from consistent data across business units and increased confidence in data quality and documentation of data processes.

According to the Data Governance Institute [12] there are seven generic goals for Data Governance:

- Enable better decision-making.

- Reduce operational friction.
- Protect the needs of data stakeholders.
- Train management and staff to adopt common approaches to data issues.
- Build standard, repeatable processes.
- Reduce costs and increase effectiveness through coordination of efforts.
- Ensure transparency of processes.

To help understand what a data governance framework should cover, DAMA define the following ten data management knowledge areas as important components in the interaction between data governance and data management [7] [34].

- **Data Architecture** - The structure of an organization's data assets and data management resources.
- **Data Security** - The practice of protecting digital information and ensuring confidentiality, access control and privacy.
- **Data Storage and Operations** - The design, implementation, and support of stored data to maximize value.
- **Data Modeling and Design** - Creating a visual representation of a software system and the data elements it contains.
- **Data Warehousing and Business Intelligence** - Collection and storage of data and the methods used to analyze this information to support decision making.
- **Data Integration and Interoperability** - The alignment between independent systems to share data between each other.
- **Metadata** - A set of data that describes and gives information about other data.
- **Reference and Master Data** - Managing shared data to reduce access redundancy and ensure better data quality through standardized definition and use of data values.

- **Documents and Content** - Storage and access to data in structured or unstructured sources available for integration and interoperability with structured data.
- **Data Quality** - How well a data set is to serve its specific purpose.

2.4 Data Management

Data management is the practice of collecting, keeping, and using data securely, efficiently and cost-effectively. Managing data in a digital organization involves a wide variety of procedures, policies, tasks and practices. The work of a data management team aim to work with the business to obtain value from the use of data and a robust data management strategy is becoming more important than ever as data becomes an economic factor of production in digital goods and services. Like oil once was, data is now an immensely untapped asset and for those whom learn how to extract it are likely to profit from its use. Today, as more businesses begin to view data as one of their core assets it is becoming clear how comprehensive of an asset data can be. It is not only the data points themselves that are valuable. Reporting, analytics, insights, and value gained from people having access to data are all part of the overall value [46].

Data Management supports this value by ensuring [55]:

- **Access to data** and the ability to harness and collect information from its sources.
- **Data Quality** - Ensuring data is accurate and useable for its desired purpose.
- **Data Preparation** - Enabling data to be ready for analytics and reporting through cleaning, moving or transformation.
- **Data Integration** - Combining various sources and types of data.
- **Data Federation** - Combining data from multiple sources without having to move and store the combined data in a new location.
- **Data Governance** - Rules and decisions that help manage data to ensure

alignment between business strategy and data strategy.

- **Master Data Management** - Managing all common and essential data to create a signle version of the truth.
- **Data Streaming** - Analyzing or visualizing moving data in real-time before it is stored.

2.4.1 Data Platform

Cognizant describes a data platform as a software solution that aggregates information in a cloud-based hub, where it can be governed, accessed and delivered to users, application and/or other technologies [9]. It does so by consolidating information from various sources and enables collaboration decision-making within the organization by combining different data. Some of the benefits of a cloud-based data platform are described as [9]:

- **Data democratization** - Data is removed from silos and can be made available across the business.
- **Competitive advantage** By standardizing both structured and unstructured data, a data platform enables businesses to compete on data as an asset.
- **Governance** - With a data platform, an organization can better manage its data governance strategy, including which data to collect and who can access it.

2.4.2 Data Products

Data products are the principle of treating data as an asset or as a product [58]. Data products are often related to domains, such as drilling, completion or intervention in a D&W team. Each domain own their data products and act as the subject matter experts for the data they generate. Optimally, they should understand what data they have available, what that data means and the context of it.

It is important that data products are available to those who need it and with the ability to help themselves to what they need it. To enable data to be treated as a

product and ensure the customer's needs have been satisfied, a data product have the following list of Zhamak's principles [19]:

- **Discoverable** - Data catalogues and metadata enable lineage, owners and source of origin to be available for the data products.
- **Addressable** - A data product should be programmatically accessible and common conventions should be developed for ease of use.
- **Trustworthy and truthful** - Data products enable ownership towards data quality and data product owners should ensure metadata is included to provide customers with data provenance and lineage.
- **Self-describing with semantics and syntax** - Data schemas with semantics and syntax to ensure self-service of acquiring data products with set requirements.
- **Interoperable and governed by global standards** - Global standards establish interoperability of an organization's data.
- **Secure and governed by global access control** - Data products must be secure. Access should be managed and assigned at a granular level to fine tune the needs of the organization.
- **Version control** - A data product should take software best practices of distributed version control so data products can be branched and tested before deployed.

2.4.3 Data Quality

Data quality measures the condition of data based on factors such as completeness, consistency, reliability, accuracy, and if it's up to date. By measuring data quality levels organizations can identify data errors that need to be resolved and assess whether the data is their IT systems is fit to serve its intended purpose [57]. Poor data quality can result in inaccurate analytics and mistaken business decisions. As a first step toward determining data quality levels, organizations typically construct data asset inventories in which the accuracy, uniqueness and validity of data are

measured. The data inventory can then be compared against the data in systems on an ongoing basis to identify new data quality issues based on data quality rules created by business requirements for both operational and analytics data [57].

2.4.4 Data Catalog

A data catalog leverages metadata and data management tools to create an inventory of data assets within an organization, allowing users to find and access information quickly and easily [30]. It utilizes metadata to describe data and to create a searchable and informative inventory of all data products or assets in an organization. These data assets can be:

- Structured data (Tabular, SQL)
- Unstructured data (Documents, web pages, email, mobile data and images)
- Reports and query results
- Machine learning models
- Connections between databases
- Data visualizations and dashboards

A data catalog will enable *data citizens* (data analysts, data scientists, data stewards, and other data professionals with access to business data) to search through the organization's available data assets and help themselves to the most appropriate data for their analytical or business purposes [30]. In addition, the data catalog can include the functionality of collecting and enriching metadata associated with the data asset. A data catalog also provides the following functionality for its users:

- Search the catalog
- Automate the discovery of potentially relevant data for which they did not specifically search
- Govern the use of the data in compliance with industry or government regulations

2.4.5 Master Data Management

Master Data Management (MDM) is focused on master or reference data and is described as the consistent and uniform set of identifiers and extended attributes that describes the core entities of the enterprise [26]. MDM pulls together multiple data items that relate to the same logical object to create a single version of the truth across for the same business object the organization. Often, there is no agreement on how common data item should be stored in different systems so when trying to combine records from different systems developers often face the problem of which source to select as the most trusted and accurate [49]. A typical example in the oil and gas industry is well header information which describes the most common attributes of a well like its name, location and type. Across all the various subsurface and top-side systems in the organization, the well header information is often misaligned which then cause issues when attempting to compare the same logical object between each system. In the ideal application/data architecture, master data is directly linked and connected with all the transactional systems, as shown in figure 2.1. Clean and accurate attribution for each master data organization entity is maintained in the same master data management repository, and the master data management system can supply these attributes back to connected transactional systems. At its core, MDM ensures that an organization does not use multiple and possibly inconsistent versions of the same master data in different parts of its operations [33].

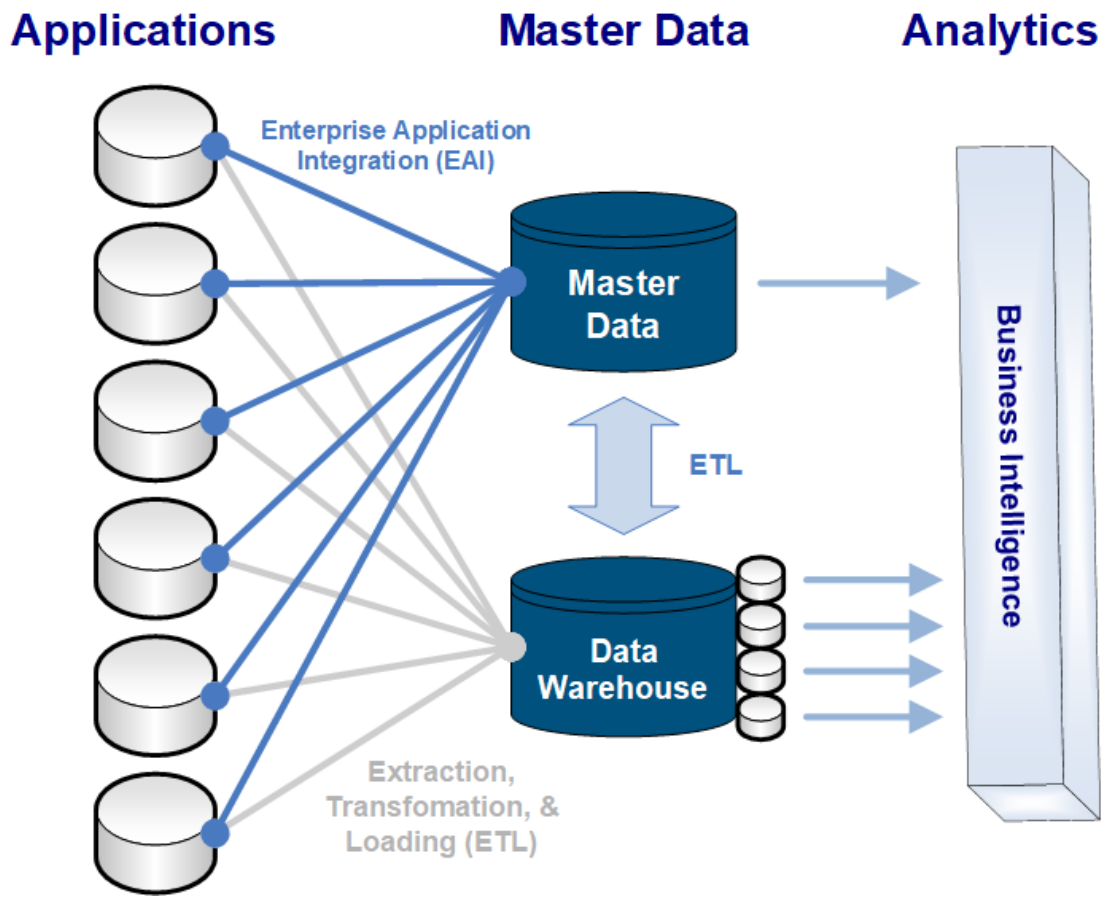


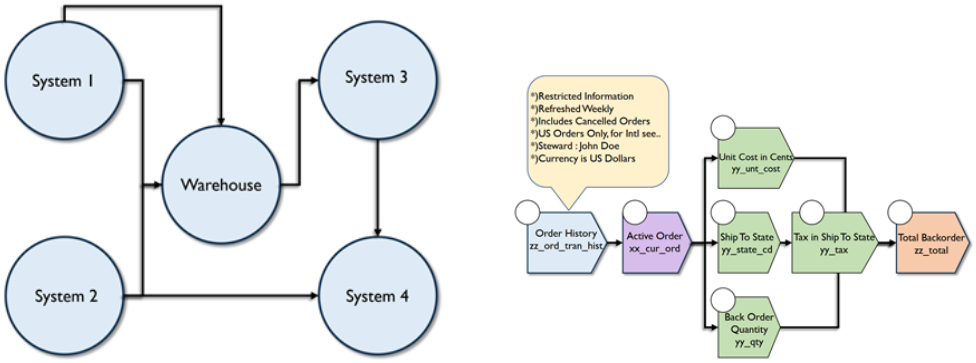
Figure 2.1: Master Data Management Architecture [33]

2.4.6 Data Lineage

Data Lineage is part of data management practices and the exponential growth of data volumes in recent years has sparked its popularity. Den Norske Dataforening [13] describes data lineage as the process of how data is created or acquired, processed and its general journey within the organisation. Data lineage is how we can ensure data quality in how we create, handle, and share data, build trust internally towards data and promote the ability to read, work with and analyse data. With data lineage, data engineers are able to identify the impact of changes they are looking to make and business analysts see where their data is coming from and ensures that the right data is used to drive business decisions [10]. The International Asso-

ciation of Data Management Professionals [34] defines data lineage as "the pathway along which data moves from its point of origin to its point of usage" and describes how metadata is a prerequisite for proper data lineage.

There are two different types of data lineage. Business lineage which aim to provide a summary view and visualize where data is coming from and to, and technical lineage which allows IT and data architects to view transformations and query-level lineage on the data pipelines on every alteration in the data [10].



(a) Business Data Lineage. (b) Technical Data Lineage.

Figure 2.2: DMBOK's two dimensions of data lineage [34].

2.4.7 Metadata

Metadata is data about data. It is documentation that describes data such as details of units of measure, survey tool details, version information or codes used in the dataset [59]. It provides the context needed in order to use data effectively, it enables search and retrieval of the data and a better understanding of the relationships between different pieces of data and how that data is used [32].

Metadata management aim to ensure data is trusted and accessible to data consumers throughout the organisation. It gives data consumers visibility into what data is available, and in conjunction with data lineage provides an ongoing record of where data originates, how it moves, how it is transformed, who accesses it and where it is stored, all of which are metadata [32].

2.5 Overview of relevant Standards

Standards such as International Organization for Standardization (ISO) and the Norwegian shelf's competitive position (NORSOK) standards provide a strong basis for the development of national and international regulations, helping save time and improve quality of services and projects. They make accessing a new market easier and ensure that the technology is tested and reliable [35], and attempt to ensure that quality, consistency, and safety and may improve performance by giving organizations the knowledge they need to optimize their operations [43]. The following section give a brief overview to some of the relevant and central standards when working with engineering assets and data analytics, management and governance:

NORSOK Z Standards

The NORSOK standards are developed by the Norwegian petroleum industry as a part of the NORSOK initiative. The purpose of NORSOK standards is to contribute to meet the NORSOK goals, which are to replace individual oil company specifications and other industry guidelines and documents for use in existing and future petroleum industry developments [45]. As the Z-Technical Information in NORSOK aim to define the technical information which shall be available for use in the operational phase, the availability and quality standards of data become relevant to this research.

ISO 8000

'Data Quality and Enterprise Master Data' ISO 8000 defines characteristics of information and data that determine its quality, and provides methods to manage, measure and improve the quality of information and data [36]. It aims to provide a framework for process-centric data quality management in order to effectively and efficiently improve data quality. In addition, ISO 8000 tackle the aspects of master data and master data management.

ISO 15926-4

'Industrial automation systems and integration, Integration of life-cycle data for process plants including oil and gas production facilities, Part 4: Initial reference data' This ISO standards specify the initial set of core reference data items which

can be used to record information about oil and gas production facilities or process plants [38].

ISO 15926-14

'Industrial automation systems and integration, Integration of life-cycle data for process plants including oil and gas production facilities, Part 14: Data model adapted for OWL2 Direct Semantics' [37]. Is a standard under development and also refered to as 'Industrial top-level ontology'. It defines general and domain independent terms to facilitate interoperability of ontologies across multiple domains. ISO 15926's primary purpose is to provide a foundation ontology to support the integration and sharing of data related to the lifecycle of an asset [50].

Chapter 3

Introduction to DataOps

3.1 DataOps

Lenny Liebmann first used the term Data Operations (DataOps) in an article featured in IBM Big Data & Analytics Hub titled '3 reasons why DataOps is essential for big data success' in 2014 [29]. Since then DataOps has risen in popularity for each year and Gartner named DataOps on the Hype Cycle for data management in 2018 [25], further increasing its adoption. DataOps can be described as a new discipline which draws inspiration from DevOps, Agile and Lean for delivering data-analytics and operationalizing data management. DataOps strives to improve collaboration among data scientists, engineers, and technologists so that every team is working in sync to use data more appropriately and in less time [41].

Gartner defines DataOps as "a collaborative data management practice focused on improving the communication, integration and automation of data flows between data managers and data consumers across an organization. The goal of DataOps is to deliver value faster by creating predictable delivery and change management of data, data models and related artifacts. DataOps uses technology to automate the design, deployment and management of data delivery with appropriate levels of governance, and it uses metadata to improve the useability and value of data in a dynamic environment." [23]

DataOps is about breaking down silos and optimizing the broad availability and usability of industrial data. It considers the importance of collaboration and the focus on developing data products and activating data for business value across an organization. In short, DataOps is not just about technology or tools, but rather the collaboration between business and data disciplines, and how an organization treats its data. Whilst investigating various definitions and articles describing DataOps it is clear that it can be approached differently. Some view it as a team oriented process and concept, others are more focused on the activities of data pipelines and automation of data management tasks, whilst others view it as a process to eliminate errors and inefficiencies in data management and reducing the risk of data quality degradation. This thesis focuses on the principle that DataOps' goal is to liberate data from its sources and deliver it to the person, application or system where it produces business value by using the appropriate tools and technology in collaboration of teams.

If we investigate the Hype Cycle for Data Management [25] data lakes, data preparation, data catalogs and metadata management solutions are all nearing the 'Trough of Disillusionment' phase and not achieved its potential yet. Many companies have found these investments to face challenges of scaling the use cases and experience that data preparation takes time and the quality may be poor [41]. In a 2020 survey of global companies, McKinsey found that the volume of new analytical features can be increased by 50 percent if the organization embeds DataOps in their practices [41]. As DataOps strives to foster collaboration among data disciplines and business users, DataOps becomes an enablement tool that can be used across the value chain, shown in figure 3.1.

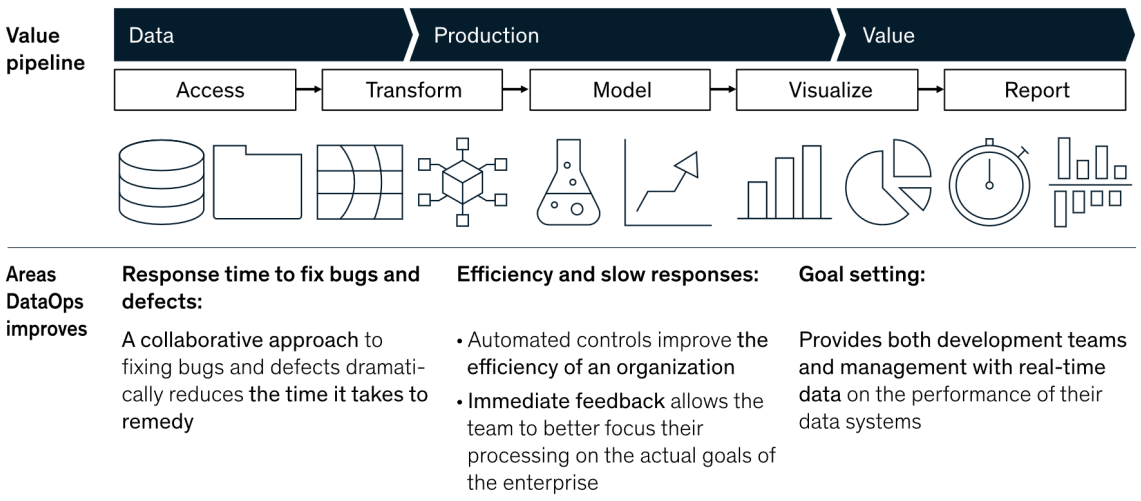


Figure 3.1: DataOps as an enablement tool across the value chain [41].

Specific benefits found in the survey was:

- **Efficient data management** - Automated data pipelines and processes enable data workers to spend more time on use case development.
- **Improved data accessibility** - By having the correct tools and processes in place to communicate the data available in the organization and breaking down data silos between systems.
- **Enterprise data governance** - DataOps enables companies to create and enforce a firm framework for managing data and providing consistency in processes with improved data quality, integration and accessibility, and ownership from the data consumers.
- **Rapid use case development** - DataOps aims to shorten the time from an analytical idea is born until it can be operationalized by working more efficiently with data throughout the organization.

Companies that embed DataOps in their organization will have to consider the key dimensions of not only technology, but also people and processes.

- **People** - A change of the skill set and culture toward the continuous usage of data and the automation of enhanced processes through data.
- **Process** - An end-to-end revision of processes to aim for streamlined and fully automated deployment of all types of new analytics models to production. To achieve a fully streamlined process, the processes of master data management and treating data as products also need to be considered.
- **Technology** - The setup of an end-to-end toolchain for full automation of the integration and deployment pipeline for models, an interface to the data platform through a data catalog and tracking data lineage of the data assets.

3.2 DataOps Principles

The DataOps Manifesto defines the following principles [39]:

1. **Continually satisfy your customer** - Through early and fast delivery of valuable analytic insights where data disciplines work in iterations with the customer.
2. **Value working analytics** - The primary measure of data analytics should be to provide analytical insight for the customer, and DataOps should deliver that by incorporating accurate data with a robust framework and system.
3. **Embrace change** by adapting to changing internal and external processes and welcome evolving customer needs.
4. **It's a team sport** - An analytical project involve a variety of roles, skills and tools. A diversity of backgrounds and opinions increase innovation and productivity.
5. **Daily interactions** - Constant communication between customer, analytic teams, and operations is required on a day-by-day basis.
6. **Self-organize** - To quickly meet the change in needs and surroundings and improve collaboration by working towards a common goal.
7. **Reduce heroism** - By creating sustainable teams and processes the group

should be able to adapt to the ever increasing needs of data analytics and not rely on individual heroism to lift the entire project.

8. **Reflect** - The team should self-reflect by having retrospectives at regular intervals to reflect on feedback, themselves and operational statistics.
9. **Analytics is code** - Fundamentally, all tools used by analytic teams to access, integrate, model or visualize data will generate code which describes the actions taken upon data to deliver the insight.
10. **Orchestrate** - A key driver for analytic success is to orchestrate and systematize the entire end-to-end pipeline.
11. **Make it reproducible** by versioning data, software configuration and code to re-use efforts.
12. **Disposable environments** - Analytic teams require disposable technical environments that reflect their production environment to minimize the efforts from development and testing.
13. **Simplicity** - Focus on the essentials needed to complete the work and pay extra attention to technical excellence and smart design.
14. **Analytics is manufacturing** - Treat your analytic pipeline as lean manufacturing lines. A key concept in DataOps is to reduce waste in the process and continuously improve the manufacturing of analytic insights.
15. **Quality is paramount** - Analytic pipelines should have a foundation capable of monitoring and managing quality and performance.
16. **Monitor quality and performance** to generate operational statistics and detect anomalies in performance, security and quality.
17. **Reuse** - Avoid repetition of previous work performed by the team or individual.
18. **Improve cycle times** - Minimize the time and effort required to turn a customer need into an analytic idea.

From the principles listed above, it is clear that the manifesto puts an emphasis

on communication and collaboration rather than tools and automation. Customer collaboration and satisfaction is of high priority and the DataOps principles focuses on having solid processes and foundations in place for how an organization gain analytic insight, and how the value lies in the collaboration between analytic teams and the customer or business.

The principles above are similar to the agile manifesto, but the agile manifesto was developed with software in mind. DataOps add some additional elements to cover the needs of a data platform.

The principles and practices in DataOps are not unique to DataOps, but a combination of proven methodologies that helped grow other industries. Agile, Lean, DevOps and Data Governance are all key concepts that play role in the definition of DataOps. Agile governs the development of analytics, DevOps optimizes delivery and verifications, lean orchestrate and monitor the data delivery and data governance aim to apply ownership and standards to data. Figure 3.2 aim to show how DataOps combines the best of each concept.

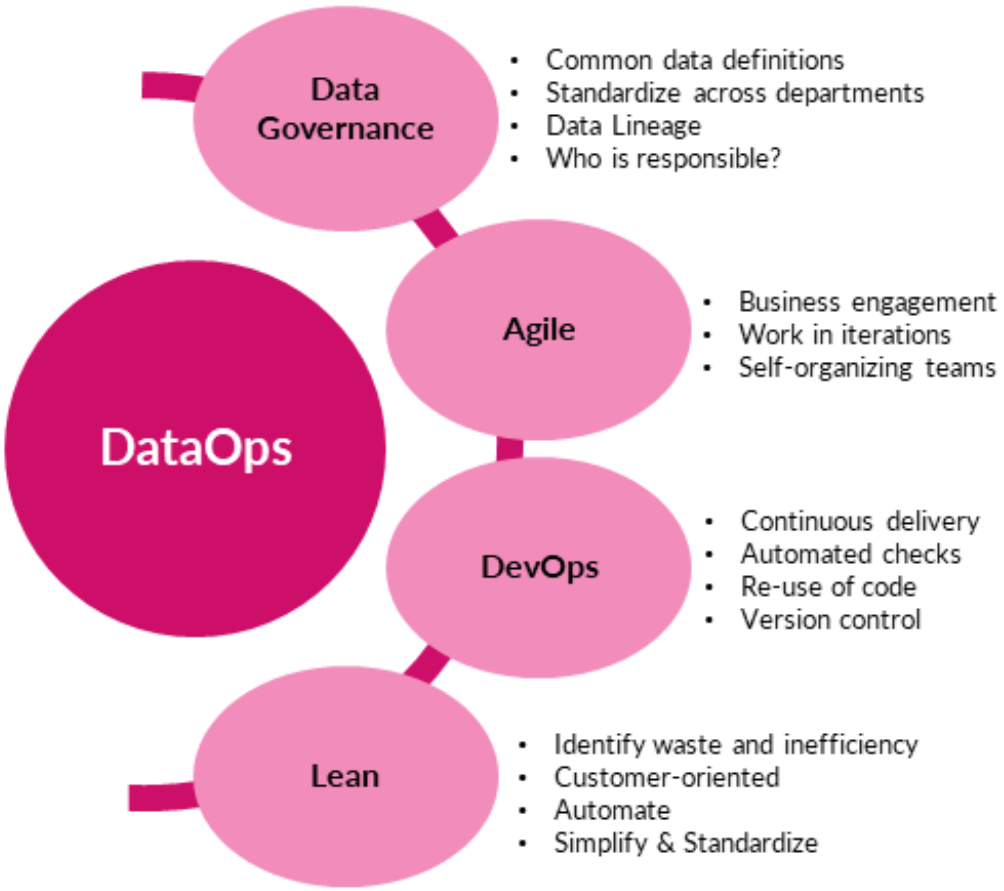


Figure 3.2: DataOps combines the best from Agile, Lean, DevOps and Data Governance.

3.3 Data Governance in DataOps

Gartner defines data governance as "the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics". As governance is concerned with policies and compliance some governance initiatives are somewhat similar to handing out parking tickets for parking violations. However, if we take a DataOps approach

to governance then the goal is not to limit users, but promote safe and controlled use of data at scale. Through active enablement of data governance through DataOps, the emphasis is on quality, management, ownership and most importantly, increased usage of data due to trust from the end-user.

3.4 Agile in DataOps

Agile project management delivers valuable features in short intervals and seeks immediate feedback. Large initiatives are broken into small increments and delivered iteratively. In Agile, the data science team responds faster and aligns more closely with the requirements and immediate priorities of end-users. The Agile methodology is particularly effective in environments where requirements are quickly evolving [6]. DataOps encourages improvement and innovation by introducing the concepts of agile development into data analytics and data management by having data teams and data consumers collaborate effectively to create effective data pipelines. The main concepts taken from agile is business engagement, working in iterations and utilizing self-organizing teams.

3.5 Lean in DataOps

Lean was originally created by Toyota to eliminate waste and inefficiency in its manufacturing operations. The process became so successful that it has been embraced in manufacturing sectors around the world. Being lean is critical in order to lower costs. The goal of lean is to eliminate waste in any process. Unless a process has gone through lean multiple times, it contains some element of waste. When done correctly, lean can create huge improvements in efficiency, cycle time, productivity, material costs, and scrap, leading to lower costs and improved competitiveness. American Society of Mechanical engineers(ASME) have defined five key principles of lean [42]. The additional descriptions give some insight into how inspiration from the lean principles is visible in DataOps.

- **Identify Value** - Creating the right concept of the Data Product definition to be used throughout the company to support governance of data assets in a

systematic way.

- **Map Value Stream** - The company wide data value streams will be made up of various data products.
- **Create Flow** - Breaking down silos and encouraging cross functional thinking.
- **Establish Pull** - Ensuring that what is produced is needed by a consumer. Tracking all usage and identifying effort wasted on unused products.
- **Seek Perfection** - Measure, identify friction and remove friction.

3.6 DevOps in DataOps

Firstly, DataOps is not just DevOps for data. Its scope is much broader and its aim more transformative [14]. DevOps is an approach to software development that accelerates the release and update of software using automation. It focused on continuous integration and continuous delivery, or CI/CD, of software by automating testing and deployment of code and utilizing infrastructure as code for on-demand IT resources. By improving the quality and cycle time of code release, release cycles for software has gone from months to seconds and enabled companies to grow at a faster pace [14]. DataOps builds upon the DevOps model of develop, build, test, deploy and run as seen in Figure 3.3 and includes how in DataOps the data pipelines ingest raw data sources as input and through a series of orchestrated steps produce analytic insight. DataOps aim to automate the orchestration and monitor the quality of data flowing through the pipelines [14]. The sandbox environment aim to have an isolated development environment where analysts or engineers can write and test new features without impact other developers.

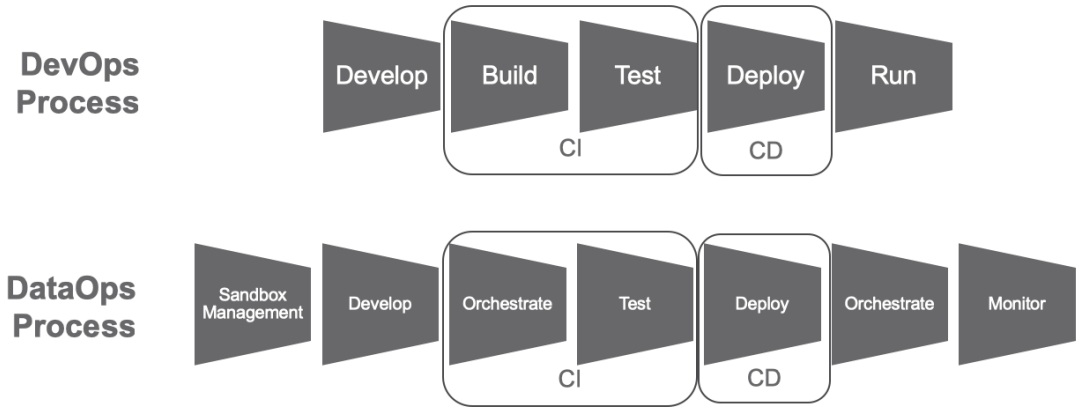


Figure 3.3: DataOps vs DevOps processes [14]

3.7 DataOps Engineer

The DataOps Engineer is responsible for the environment in which data development takes place. They are building the tools and framework that data managers, analysts and engineers are using within their workflows. Part of DataOps is applying automation to streamline workflows [16], examples would be to replace manual procedures that execute data operations with automated orchestration, monitor data pipelines to catch errors before the customers see them and construct dashboard on the data lifecycle from data retrieval and preparation to analysis and reporting.

The DataOps engineer focuses on a set of practices focused on collaboration and automation and is emerging to replace traditional data management and help drive data-intensive initiatives. For DataOps to be successful, collaboration is key and the business leaders, data disciplines and IT must work together [51]. The DataOps engineer aim to drive the data organization to higher levels of productivity and ensure value outtake from the business. Hence, one of the most important aspects of DataOps is understanding the business function. Being able to bridge the value pipelines from data to customer, and establishing processes that benefit both functions is key [16].

The investment into DataOps impacts the whole pipeline of data delivery and the

addition of a dedicated DataOps engineer to data disciplines enables data managers, engineers, scientists, analysts and self-service users to spend more time creating value and less time chasing errors and managing issues. An example of how the time allocated to errors and operational tasks, new features and data for customers and improvements change with the addition of a DataOps engineer is shown in Figure 3.4.

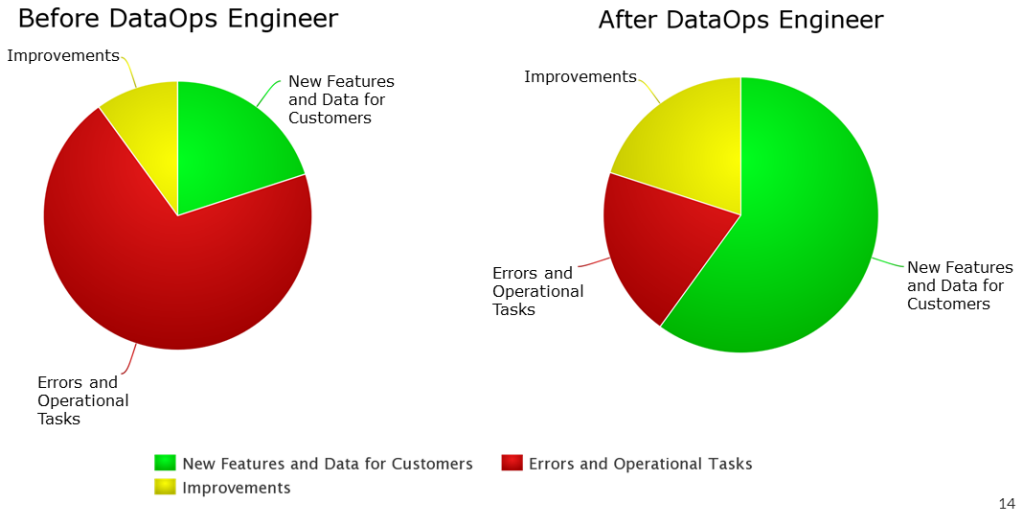


Figure 3.4: Time allocation before and after DataOps, inspired by [16]

3.8 DataOps Maturity Model

Based on the DataOps principles and concepts detailed in this chapter, discussions with Aker BP and research conducted a DataOps maturity model was developed as part of the thesis paper. One of the main objectives in the following chapters will be to establish Aker BP's current maturity based on the analysis and based on the maturity model propose measures to be taken by Aker BP to raise their maturity.

In general terms, a maturity model assess the current effectiveness of an organization and supports figuring out what capabilities should be acquired next in order to improve their performance. To define any company's maturity in adopting a DataOps framework and working with data, the DataOps maturity model was created as mentioned above. The maturity levels of the model are inspired by the five levels of the Capability Maturity Model Integration (CMMI) project work presented in [56] from the Software Engineering Institute, Carnegie Mellon University with slight iterations. The maturity levels are measured by the achievement of the specific and generic goals that apply to each process area and the maturity level details are designated by the numbers 1 through 5. The process areas for the DataOps model were assigned based on the main research topics in this thesis of Data Management and Analytics, Culture and Collaboration, and Data Governance. The following sections describe the characteristics of each maturity level in detail.

Maturity Level 1: Initial

At maturity level 1, processes are usually ad-hoc and unorganized. The organization does not provide a stable environment nor a business unit working on improving how the organization utilizes its data. The use of data is based on individual efforts and competence and not on the use of proven processes. The main purpose of gathering data is reporting purposes, and governance and data quality measures lack consistency. Organizations that are characterized as maturity level 1 has a tendency over commit and abandon processes in the time of crisis. As a result, these organizations are often unable to repeat their past successes due to lacking consistency in collaboration and delivery between data disciplines and business users.

Maturity Level 2: Develop

At maturity level 2, the organization has ensured that there are certain requirements and processes in place to perform and develop analytical insight. The value of data and its use cases is being communicated and some processes are in place to protect data quality across the organization. Data and analytics are being actively utilized in decision making and there is a company-wide strategy in place to ensure multidisciplinary collaboration.

Maturity Level 3: Define

At maturity level 3, an organization has achieved all the specific and generic goals of the process areas at maturity level 1, 2 and 3. Best practices, standards and practices are now established consistently across the organization. Dedicated teams are focused on analytical value delivery and improvement of current processes in the business. Business users are enabled by having cleaned data available from its sources or through an accessible data platform, and work in close collaboration with data disciplines. Compared to maturity level 2, the processes are now managed more proactively using an understanding of the process activities, work products and services.

Maturity Level 4: Measure

At maturity level 4, there is a universal confidence in the data and resulting insights in the organization. Business users are able to maximize business outcomes based on data analysis, and analytical insight and performance is measured and continuously improved. Certain data management practices, like data loading or quality checking, are now automated across the organization and both tools and data products are clearly defined and optimized in a catalog or similar.

Maturity Level 5: Optimize

At maturity level 5, an organization has achieved all the previous specific and generic goals and processes are now continually improved based on a quantitative understanding of the common cause of lacks and inefficiencies in the processes. Lifted capabilities are aligned with the business values and objectives of the organization and the organization's ability to rapidly respond to changes and opportunities is enhanced by finding ways to accelerate and share learning. New technology and an-

alytical insight is developed upon and architecture designed for efficiency and large volumes of data. The business is constantly looking to leverage and combine data from new non-obvious sources and data pipelines continuously improve business objectives.

The DataOps Maturity Model

The model, as seen in figure 3.5, is developed on the core principles from the DataOps Manifesto [39] as described above. By taking into consideration how DataOps is related to agile, DevOps, and lean manufacturing principles and its close relation data management and data analytics practices, a DataOps maturity model has been developed to analyse an organizations DataOps readiness.

DataOps Maturity Model

Stage	1. Initial	2. Develop	3. Define	4. Measure	5. Optimize
Data Management & Analytics	• Ad-hoc ways of working. • Few datasets or products have been onboarded to the organization. • Data is used mainly for reporting purposes.	• Data insights are used to inform business decisions. • The organization can combine data from different sources. • Some established ways of working. • Data to be defined as an important asset.	• Dedicated teams focused on analytical value delivery and improvement. • DataOps practices are characterized for the organization and proactive. • Defined data catalogs used to manage data inventory and products.	• Data delivery and insight is measured and controlled. • Analytical insight and performance is measured and continuously improved. • Certain data management practices are automated across the organization.	• DataOps processes are fully automated, regularly improved and optimized. • New technology and analytics is developed upon an architecture designed for efficiency and large volumes of data.
Culture & Collaboration	• The use of data is based on individual efforts. • The business lacks a coherent data architecture and strategy. • No business unit is working on improving how the organization works with data.	• Business goals are known and communicated. • Activities are deliberate and documented. • Data is part of measuring results.	• DataOps practices are aligned with an enterprise-wide strategy. • Business users have cleaned data available to enable data-driven insight. • Deliberate ways of collaborating across business units.	• Decision makers are enabled with the results of data analysis to maximize business outcomes. • Service catalog of agreed technologies and high-level framework for teams to adopt new technologies.	• The business is constantly looking to leverage and combine data from new sources. • Automated AI/ML algorithms and data pipelines continuously improve business objectives.
Data Governance	• Governance is largely manual and lacks consistency.	• Processes are in place to protect data quality across the organization.	• Best practices are defined and evangelized, and teams enabled to evolve practices.	• There is universal confidence in the data and resulting insights.	• Data Governance is integrated into all business processes.

Figure 3.5: DataOps Maturity Model.

Chapter 4

Review of the Current Practice in Aker BP

In Aker BP today, data management efforts are limited and based on ad-hoc need and fast changing priorities. Execution is highly manual as sufficient structure and tool support is missing. The pressure on the data management function to verify and clean up data sets is becoming larger each year as more digitalization projects are realised through EurekaX, Aker BP's digital lab. In addition, Aker BP have, together with Accenture, initiated a Data Factory to accelerate its goal of digitalizing the full lifecycle of its operations to cut costs, improve productivity, and lower its carbon footprint [1]. In 2017 Aker BP established a partnership with Cognite to develop Cognite Data Fusion (CDF) with the goal of putting raw data into real-world industrial context in a common data platform, enabling rapid application and solution creation at scale as you have data easily accessible through CDF. However, whilst both Accenture and Cognite will play an important role in achieving efficient handling of quality data, an investment into data management from Aker BP is required to allow for automation over time and ensure the right competence within Aker BP. Overall, the key challenge today is consistent implementation of data management and a lack of structure for decision-making. This causes an inconsistency in decision-making around data which in turn limits the creation of building end to-end data products [4]. This section will aim to describe how Aker BP's data manage-

ment team functions today, how it interfaces towards the other business units and what tools and processes are being utilized.

4.1 Digital Strategy

4.1.1 Digital Transformation, EurekaX

EurekaX is the digital development program in Aker BP. The program provides the framework for running all digital development initiatives, from allocation of budget through to operational implementation. EurekaX is an investment umbrella, it is a development muscle and it is a competence center and it is changing and adapting to deliver on their core assignment: to make sure all digital initiatives are best positioned to deliver value to the business.

There are today seven prioritized end-to-end (E2E) processes and a Corporate portfolio which together make up the eight value streams. Each with a dedicated SVP from the business which is accountable for their success. Each value stream has a set of business capabilities, which again has a set of strategic themes. Capabilities and streams are fully evaluated on a yearly basis and adjusted continuously through the year.

Eureka started with a strong engagement from owners and top management, Aker BP established its digital program in July 2017. Data liberation was identified as key for a successful realization of their digitalization vision. After carefully considering the options in the market, Aker BP established a strategic partnership with Cognite, a software start-up, to develop the Cognite Data Fusion (CDF) as a foundation for the digital journey. During the first year of the program, the partnership with Cognite sparked a lot of use case ideas and engagement. At the most, Aker BP worked on over 300 use cases in different maturity stages and both parts learned a lot about combining O&G domain expertise with software expertise. This approach was good for learning, but not effectively driving progress and value in the use cases.

In 2018 it was decided to review the way Aker BP worked on the use cases, and most importantly, what use cases they worked on. They called this phase "Boozt" and the outcome was a prioritized set of use cases, organized in a new program

set-up where they would use agile working methodologies to more efficiently drive progress. The Eureka set-up and scope for the first six months was approved by top management before the summer in 2018.

In the Eureka program setup, use cases were organized into crews with fully dedicated, co-located and cross-functional teams with a clear performance contract and customer in the line business. The Eureka program sparked enthusiasm for digitalization and proved that Aker BP were able to develop minimum viable products using agile working methodologies. For each year, the program has developed and the scope of work expanded. Many learnings and experiences were made from 2017 till today and has laid the foundation for the new EurekaX setup which targets stronger customer relations, unifying the company's digital efforts and securing value outtake from digital initiatives. The EurekaX crews are identified as one of the key customers of the data factory and new DataOps function.

4.1.2 Data Factory

In 2021, Aker BP launched the Data Factory together with Accenture and Cognite to accelerate its goal of digitalizing the full lifecycle of its operations. Cognite's goal was to apply its Cognite Data Fusion software to liberate and contextualize data across information and operational-technology siloes. Whilst Accenture was given the mandate to deliver the project team that will deliver more reliable data by applying automation, innovation, and technology. The main objective of the Data Factory is to deliver sustainable data products of known quality, trusted and managed through their lifecycle, to enable Aker BP's digital transformation and corresponding capabilities as defined by the business. Data products in this context are sets of data that are ready for consumption by data consumers and was prioritized as Aker BP has multiple sources of the truth across different applications and data stores that are not interoperable resulting in duplication of data, inefficient workflows, poor data accessibility and lack of trust in the data. In addition, for the scenarios where data exists in one source, data quality and governance is not consistently applied and measured across the organisation.

The DataOps organization's vision, as described previously, is in line with the goal

of the Data Factory, and organizationally the Data Factory is owned by EurekaX. In collaboration with data management personnel, EurekaX crews and Subject Matter Expert (SME)s from the business, Aker BP work with Accenture to define the foundation for data products required to support the business deliveries and digital products. These data requirements are designated to the product team within the factory who works together with the crews, business and product owners to analyze, understand and refine the requirement into a data product. The product team breaks down input to features, that are manageable pieces of work that can be completed in a product increment of 10-12 weeks. The Chief Product Owner, Aker BP's owner of the factory, is responsible to prioritize the data product backlog. Further, the squads within the factory will take prioritized features from the backlog and break these down further into user stories that can be implemented in two-week sprints. They will work on the most valuable item in the backlog, implementing the feature user story by user story. Before an user story can be closed it needs to meet the quality standards and the acceptance criteria provided by the Product Owner (PO), which should be an Aker BP representative. Examples of outputs from the factory can be legacy data cleaned up according to new data quality standards or a data product readily available in CDF or from its source system through an Application Programming Interface (API) developed by the factory.

4.2 Culture & Collaboration

The digital transformation in Aker BP is more than technology. Aker BP aim to build digital capabilities, develop digital mind sets and implement new ways of working where decisions are driven by data. A digital agenda as Aker BP's needs to be backed by an understanding of data and knowledge of different technological concepts. To reach that objective and become a data driven company Aker BP have decided they require clear digital job families and digital roles, therefore a project currently being run to map out the required roles in a digital Aker BP. In addition, a new business unit Digital Business Transformation (DBT), previously Information Management & Technology (IM&T), was established as of january 1st 2022 and will act as the organizational home for technology and digital roles in the company. The vision of DBT is to "Enable creating the leading independent offshore E&P

company" and this will be achieved through the five functions of:

- **Supply Chain Management and Logistics & Marine** - Focuses on increasing flow efficiency and streamlining logistics.
- **Corporate IT & Cyber Security** - Provides industry leading performance and competence within cyber security, IT management and operations of digital services and technology.
- **DataOps** - Establishing a strong enabling function for data and information management.
- **Enterprise Architecture** - Connecting company strategy and business operating model with technology opportunities.
- **Digital Transformation (EurekaX)** - Fully transform core end-to-end work processes enabled by data, ecosystem, and people.

Figure 4.1: Digital & Business Transformation Organizational Structure.

The DataOps function will provide a competence home for everyone working within the DataOps disciplines across Aker BP as well as a much stronger governance and customer-centric capability to drive maturity and prioritization of demand. However, Aker BP has just embarked on the DataOps journey and whilst they are starting with fresh eyes, it is important to consider past experiences. This chapter aims to evaluate the current practices in Aker BP today, investigate what is available in the market and what Aker BP need to consider on their journey to mature their DataOps organization.

As seen above, DataOps as well as Digital Transformation is part of the business unit Digital Business Transformation (DBT). The other business units are the likes of Drilling & Wells (D&W), Exploration & Reservoir Development (ExpRes), Projects (PRO) and Operations & Asset Development (OAD). This section will focus on how the DataOps function is organized internally and its interface with Digital Transformation in DBT and D&W, also referred to as the business.

4.2.1 DataOps

DataOps has the mandate to establish an organizational structure, a governance framework, and an information architecture for data collection, management, and distribution. The DataOps function will operationally ensure that this happens in a safe, user-friendly, and cost-effective way in collaboration with product and process owners and be in line with company development. Figure 4.2 reflects that DataOps should be closely related to their internal customers, such as D&W, and strengthen their ability to cooperate across business units. In addition, it is a governing body which together facilitates insights regarding business needs and have enablers that deliver and support prioritized services. As of the writing of this thesis the DataOps organizational chart has not been established. As such, Figure 4.2 show the teams in the previous organizational structure who will likely be part of the DataOps function.

Figure 4.2: DataOps Organizational Chart

- **Data Foundation** acknowledges and understands how data requirements are based on domain expertise and business needs. They ensure that DataOps has a total portfolio overview and supports the business domains in their maturity

journey with subject matter experts and close collaboration with the data factory. The data foundation ensures a holistic approach to Data Management within and across value chains.

- **Overground and Underground Data Management** hosts experts and several Centers of excellence and provides services and capacity within information & data management, data science and analysis services.
- **Insights, Availability & Visualization** develops and manages data platforms and data products, ensures efficient data flow and distribution and considers strategy and system ownership. Examples of this would be CDF, OSDU, Data Warehouse and deliveries from the Data Factory.
- **Geographic Information System (GIS)** deliver data and insight to facilitate decision-making and reduce risk through a one-stop web solution for geographical information and infrastructure data.

4.2.2 Drilling & Wells

Aker BP's D&W business unit's mission is to deliver safe and efficient exploration and development wells to sustain Aker BP's strategy for execution and growth. The goal is to manage the drilling and well operations in full compliance with company management system and authorities, rules and regulations always respecting the environment and the people.

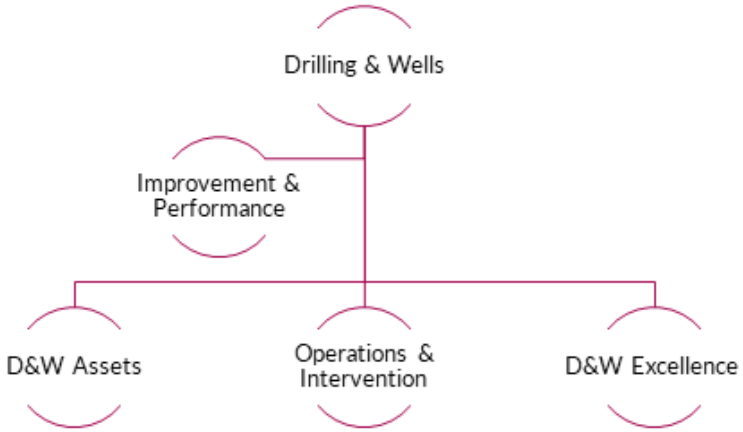


Figure 4.3: Drilling & Wells Organizational Chart.

To reach Aker BP's vision of becoming the leading independent E&P company, D&W aim to maximize value creation for the asset portfolio throughout the well lifecycle. To achieve this vision, five enablers have been identified describing how the business unit works as shown in Figure 4.4.

Figure 4.4: Drilling & Wells Enablers.

Technology and digitalization is acknowledged as one of the five enablers for D&W and will be important to achieve reduced cost of operations and increased value

output through improved decision making and efficiency. The technology and digitalization enabler includes the following drivers:

- **Technology** - Gradually improved access to reserves with lower cost and carbon through innovation and technology.
- **Improve data sources** - Technology such as sensors, smart systems and wired pipe will provide more and higher quality data.
- **Manage data** - Solid data management needs to be in place to ensure data integrity and control.
- **Liberates data** - Data and open source systems are available
- **Intuitive data** - User friendly experience and APIs for applications

The potential output if successful with this enabler will be availability of quality data, acceleration of day-to-day work, enhanced quality and consistency in deliveries, new insights from the liberated data, and improved decision making and increased planning efficiency and execution performance. In order for D&W to succeed with the enabler they have identified that data is the responsibility of all, one must embrace new tools and ways of working, and contribute with improvement ideas.

4.2.3 Summary

Today, the function of underground data management in DataOps and D&W collaborate closely and there is a decentralized data teams sitting with the business. The data management team supports the business and are working closely with them, but not organized within their teams and do not attend their team meetings. It is important to note that there is currently an Improvement & Performance (I&P) team within D&W today, this unit is working in close collaboration with the business users to deliver data analytics, performance data and drive improvement initiatives within the BU, similar to the ambition of the new DataOps function. As mentioned, Underground Data Management is located outside of D&W and as part of DBT and the I&P team within D&W, however from discussions with both key DataOps stakeholders and D&W I&P stakeholders it is clear that there are similarities in the objectives of each team and potentially conflicting future deliverables.

The participants are positive to the change and have an ambition that DataOps will deliver higher quality data to the I&P teams such as their analysis also improves, in addition to advising on best practices and collaborate on projects.

4.3 Data & Analytics

As described in earlier section, Aker BP has an ambition of being a data-driven company. A data-driven company embraces the use of data in decision making and treats data as an strategic asset of the company by making it widely available and accessible. It focuses on capturing, cleaning, and curating meaningful data from across the business. Being successful with a data-driven culture requires all levels of the organization to understand that a strong foundation of data is critical for differentiating the business from its competitors, and a belief that data helps everyone perform better.

However, albeit a focus on developing technology and becoming a digitized company, interviews and discussions conducted with personnel highlight unrealized potential of working with data in the organization. The following freely-paraphrased quotes are some of the input captured: "We want to be able analyze data faster, differently, and we want to get more out of our data. A transformation of the way we work with data and manage data, is critical to succeed with the digital transformation."

"More than 50% of time goes to finding and validating if the data is correct or not. We maybe use 5-10% of the data available."

"We are really struggling to maximize the utilisation of data - and minimizing the uncertainty in our data."

The statements given from business users above show the lack of trust in data and how Aker BP is still unable to utilize the data they have available.

4.3.1 Data Quality

The information given below is based on multiple data managers, with 5-10 years of experience in Aker BP, analysis of the data quality in the key D&W systems. In the

O&G industry in Norway there are multiple regulations in order to be an operation on the NCS. Those regulations include reporting requirements and as such Aker BP have processes in place to capture the necessary data. However, business users are still struggling to find and validate if the data is correct or not. In many of the research discussions business users explain how they struggle when comparing multiple sources of the same data as it is often misaligned. This raises the need for clearer systems of record and governed data products where the data quality has already been confirmed. Today data quality checks may be performed multiple times as no processes are in place to execute and document data management tasks. Few best practices are created, no overview of data products with their corresponding system of record and quality for each asset and unclear alignment with the business for their data requirements are issues highlighted during interviews and data quality suffer as a result.

4.3.2 Liberating and working with data

In Aker BP today, there is no common data management tool available. Data managers rely on database queries, APIs or the application's GUI to handle data and efforts are based on individual's expertise to efficiently process data and verify quality. The data managers today struggle with being occupied in ad-hoc data deliveries or cleaning and operational day-to-day tasks, and there has been little time for improvement of data management tasks. A data management tool could enable a stricter practice of monitoring, enforcing data requirement and aid in defining data quality and cleaning standards in addition to breaking down the data silos.

To enable Aker BP's goal of liberating data and becoming a data-driven company, there are already tools in place that are breaking down the data silos. To enable self-service and quick delivery of insight to the business users from the data disciplines, Microsoft's Power Platform and specifically Power BI is being heavily utilized in Aker BP. One of the key benefits is in data federation and being able to combine multiple data sources into one analytical dashboard through Power BI. Power BI is enabling Aker BP to easily achieve value from both streaming and reported data from their operations and can be used for further analysis in planning. In addition to PowerBI, CDF and OSDU has been pointed out as key initiatives going forward.

Cognite Data Fusion

Cognite Data Fusion (CDF) aim to make data available, readable and open for connections to other relevant data. For Aker BP the solution contextualizes data as a service and supports Aker BP's vision to liberate data from silos and bringing it together in one place. Further transforming data into useful information for better decision making and innovation. Industrial applications in the O&G industry are so complex which has led to data being stuck in their respective application. For Aker BP vendors and internal project point-to-point connection to each system has had to be set up which is time consuming and costly to maintain. The goal for CDF was to enable any consumer to connect and have the necessary data easily available. However, whilst the vision is clear and Aker BP has been able to liberate multiple data products through Cognite's platform, it is clear from discussions with stakeholders that Aker BP has not reached CDFs full potential and several data types are missing still. From a D&W perspective only a few data types are ingested as of today, and being able to connect additional sources will be of huge benefit to Aker BP if successful.

Open Subsurface Data Universe

Open Subsurface Data Universe (OSDU) aim to establish an open source data platform for cloud-native datacentric applications to enable access to the full range of subsurface and wells data, as well as supporting existing applications and data frameworks [47]. The data platform will separate data from applications and standardize data definitions and APIs to reduce silos and enable a data-driven subsurface community for operators. Both business users and data managers in Aker BP believe OSDU can be a key enabler for a seamless float of data across applications in their underground portfolio. Whilst still in an early stage, it is clear that OSDU has both operators and vendors are now working together in solving one of the largest data related issues in eliminating data silos. The biggest challenge may be in being able to communicate the strategy for OSDU within Aker BP, as from discussions it seems unclear for most what it means in practice and how OSDU will interface with CDF which seemingly has a similar vision in terms of eliminating data silos.

Power Platform

Power Platform combines Power Apps, Power BI, Power Automate, and Power Virtual Agents into a single business application platform in order to provide app building and data insights. Aker BP utilize all four technologies, whilst Power BI has become the most used application for liberating data it is clear that it lacks governance and can be improved.

From analysing the as-is workflows in Aker BP it is clear that the most popular tool for delivering data analytics and insight is Microsoft's business intelligence tool Power BI. Power Bi aim to connect to and visualise any data from your data sources with the scale to analyse, share, and promote insights across the organisation while maintaining data accuracy, consistency, and security [44]. One of the largest benefits of Power BI is that the platform empowers both self-service users and data analysts in the organisation. In Aker BP the use of Power BI has clear benefits, but there are also some weaknesses and lack of governance in the current utilization of the tool. Figure 4.5 shows a Power BI Data Governance Framework, Aker BP utilize a variation of the 'IT-managed self service BI' which focuses on having established data flows from the data management teams, or data sets and models from a data warehouse. The data is owned by the business, maintained and made available from data management, and reports and dashboards are created by self-service users and I&P teams in the organization. The Power BI dashboards can be handed over and maintained by a group in the IT organization, however due to the lack of governance and common work processes, most dashboards in Aker BP are today maintained by those who made it. In addition, there are no clear governance or overview on what data sources are available and to what standards a report should be made. From the discussions with business users it is very evident that the use of Power BI has become person-dependent as they rely on a few key members of their teams to develop and maintain the dashboards and as such the dashboards created see a rapid decline in usage as the key members move into new roles. This had led to lack of trust in the dashboards available, and multiple dashboards with the same functionality has been created without removing the redundant ones. Overall, it is clear that Aker BP need to establish better governance in the delivery model between DataOps, I&P and business users in order to continue the value outtake that comes from having

data and analysis made easily available through Power BI.

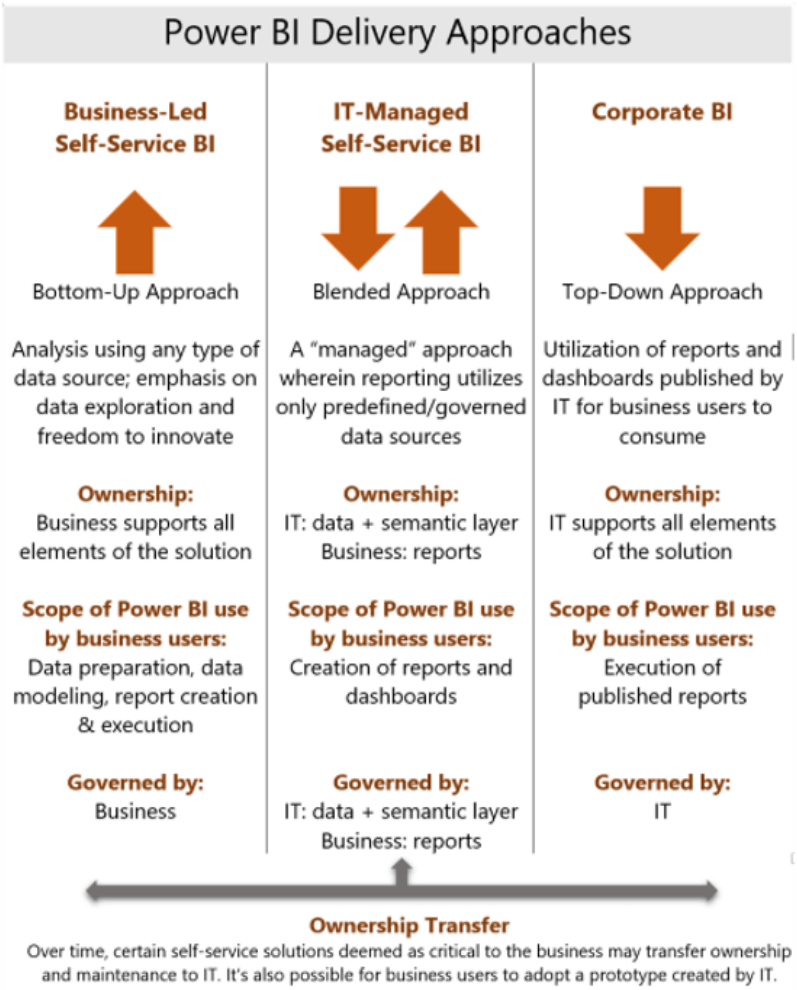


Figure 4.5: Example of a Power BI Delivery Approach or Governance Framework [5].

Dr. Q

Data Reliability & Quality (Dr.Q) is an application developed Petrotechnical Data Systems (PDS) in collaboration with Aker BP to validate real-time Well-Site Information Transfer Standard Markup Language (WITSML) data during operations where data is captured and provided from WITSML. Dr.Q greatly simplifies the management of data quality during drilling operations, highlighting and escalating data issues and helping to quickly identify their root cause.

The system calculates Problem Time, which is the basis for a set of KPI reports used to support contractual bonus/penalty conditions in Service Provider contracts. The current set of validation rules monitors WITSML wells, wellbores, logs, curves and trajectories.

The Dr.Q application delivers business value by making it simpler and faster to:

- Manage a common mnemonics catalog
- Manage curve ordering to Service Providers
- Detect real-time data-related issues,
- Identify the source (and in some cases cause) of the issue,
- Verify when the issue has been resolved, and provide a continuous overview of the summary health of the service.

Better insights into the quality of data and faster intervention results in a higher confidence for downstream consumers of the drilling data that the data is correct. As operators continue to build upon their integrated drilling data services, Dr.Q is helping to ensure the critical supporting data is complete, up to date and valid. As data-driven operational support becomes more prevalent in the industry, a tool such as Dr.Q is essential to establishing and maintaining an accurate, reliable data platform to serve downstream applications and analytics.

4.4 Data Governance

At its foundation, Data Governance aim to enable an organization to provide trustworthy data through defining roles and responsibilities, policies and processes in the organization. Key findings from the Data Foundation blueprint [4] and discussions with business users and data managers show that in addition to data management functions, Aker BP also need to mature its data governance function to reduce inefficiencies caused by lack of trust in data. The goal with data governance in Aker BP is to ensure that data and related capabilities are clearly governed and maintained over time. As discussed above, the lack of governance for the organizations Power BI strategy and dashboards are one of the clear examples of why Aker BP is now

placing data governance as one of the top priorities in DataOps. Information must have clearly defined ownership, governance and stewardship throughout its lifecycle.

Aker BP D&W have established an application list with defined 'Information Definition Owners' who should ensure that definitions, standards and requirements are defined, documented and communicated and define, monitor and maintain information integrity and ensure information quality. 'System Owners' who has ownership of applications within their area of expertise and work processes as defined the in the Business Management System (BMS). Finally, 'Super Users' that support the peer group to better use application and to improve workflows and to test any changes to the given application.

The 'Information Definition Owners' are, as mentioned, responsible for the information and data quality in the application or system. However, in many cases it is clear that whilst the role is defined it is not carried out to its description. Data managers in Aker BP explain how it is, in some scenarios, difficult for them to convince the business or the information definition owner that the poor data quality in their system is a problem. From the researchers observations this ties into the fact that data is in many systems mainly used for reporting requirements, and the business value from utilizing the data is missed. One experiment conducted during this thesis was to create a simple Power BI dashboard to track a data type for one of the D&W operations. The dashboard used one of the already established data flows from one of the key reporting data sources during a drilling event and aimed to visualize one of the required data types for reporting. Instead of just visualizing the data, the dashboard could be used to track Key Performance Indicators (KPI)s during drilling by the business users. The result was a more conscious relation to the data quality as the output of the Power BI dashboard looked off to them if there was data missing. Ownership towards data seem to be one of the areas where Aker BP can improve, and creating policies, processes and guidelines to follow to ensure good data quality may prove inefficient. This will be further elaborated on in the coming chapters.

4.5 Summary

Today, Aker BP's data management efforts are heavily focused on applications and data is managed as part of the application by data managers. Due to capacity and high workload, not all applications have any data management, and as such the data quality suffer. For the applications with data management the data quality is better, but not without issues and legacy data is often of poor quality. In addition, some of these applications have been in the industry for decades and are not up-to-date with today's requirements for data availability and flexibility. As such, data is not easily shared across systems and not being able to extract data from key systems seem worrying.

Data insights are used to inform business decisions as seen by the focus on new digital solutions and data is treated as one of the most important enablers for the business. Data is being treated as a product and there are many functions working on achieving the most out of the organizations data in order to become more data-driven. There are some established ways of working in relation to delivering data analytics and managing data through a close collaboration between data management, I&P and the business users. In short, data management aim provide the good quality data in data pipelines to I&P so that they are able to focus on delivering data analytics to the end-users. In practice, there are still ad-hoc ways of working as these processes are person-dependent and not well defined and optimized. Data is actively used in the organization, and DBT with the new DataOps function aim to further put evangelize and evolve best practices when working with data and analytics. The business goal of Aker BP is well communicated and it is clear that both technology and data will be vital in Aker BP's further advancements. Overall, the current state of Aker BP indicate a stage 2 'Develop' level of maturity in the DataOps Maturity Model in Figure 4.6. Chapter 6 will address some of the necessary steps to further mature their DataOps function.

DataOps Maturity Model

Stage	1. Initial	2. Develop	3. Define	4. Measure	5. Optimize
Data Management & Analytics	• Ad-hoc ways of working. • Few datasets or products have been onboarded to the organization. • Data is used mainly for reporting purposes.	• Data insights are used to inform business decisions. • The organization can combine data from different sources. • Some established ways of working. • Data to be defined as an important asset.	• Dedicated teams focused on analytical value delivery and improvement. • DataOps practices are characterized for the organization and proactive. • Defined data catalogs used to manage data inventory and products.	• Data delivery and insight is measured and controlled. • Analytical insight and performance is measured and continuously improved. • Certain data management practices are automated across the organization.	• DataOps processes are fully automated, regularly improved and optimized. • New technology and analytics is developed upon an architecture designed for efficiency and large volumes of data.
Culture & Collaboration	• The use of data is based on individual efforts. • The business lacks a coherent data architecture and strategy. • No business unit is working on improving how the organization works with data.	• Business goals are known and communicated. • Activities are deliberate and documented. • Data is part of measuring results.	• DataOps practices are aligned with an enterprise-wide strategy. • Business users have cleaned data available to enable data-driven insight. • Deliberate ways of collaborating across business units.	• Decision makers are enabled with the results of data analysis to maximize business outcomes. • Service catalog of agreed technologies and high-level framework for teams to adopt new technologies.	• The business is constantly looking to leverage and combine data from new sources. • Automated AI/ML algorithms and data pipelines continuously improve business objectives.
Data Governance	• Governance is largely manual and lacks consistency.	• Processes are in place to protect data quality across the organization.	• Best practices are defined and evangelized, and teams enabled to evolve practices.	• There is universal confidence in the data and resulting insights.	• Data Governance is integrated into all business processes.

Figure 4.6: DataOps Maturity Model - Aker BP's maturity in red based on the analysis conducted.

Chapter 5

Introduction to State of the Art in the Industry

Data Management tools are used to develop and monitor practices, as well as organize, process, and analyze an organization's data. These tools are designed to arrange and harmonize data, and the aim is to provide a high degree of efficiency and effectiveness when working with data. Data Management tools also support privacy, security, and the elimination of data redundancy. Effective Data Management uses a combination of software tools and best practices to control and organize data resources effectively [18]. DataOps tools brings together several kinds of data management practices such as data ingestion, data transformation, data analysis, and data visualization in an attempt to automate and simplify all data life cycle phases. A DataOps platform can be used to unify all the development and operations of data workflows and support a DataOps mindset towards collaborating on data analytics. This chapter aim to cover some of the data management and DataOps tools to be considered for Aker BP.

5.1 Overview

DataOps tools should serve as command centers for DataOps teams. These platforms orchestrate individuals, processes, and technology to provide a reliable data

pipeline to their consumers [52]. In addition to a DataOps tool for data pipelines, Aker BP are also lacking a common data management platform to clean, validate, and map data in Aker BP. Key data management personnel has stated that if Aker BP wants to pull all data classified as Aker BP internal, they are not able to do so because their systems are incapable. To add to that, there is no common overview or catalog, nor common governance, of data products or data types in the company. Aker BP is in need of an effective platform to support data management activities and meet the business needs of more available data.

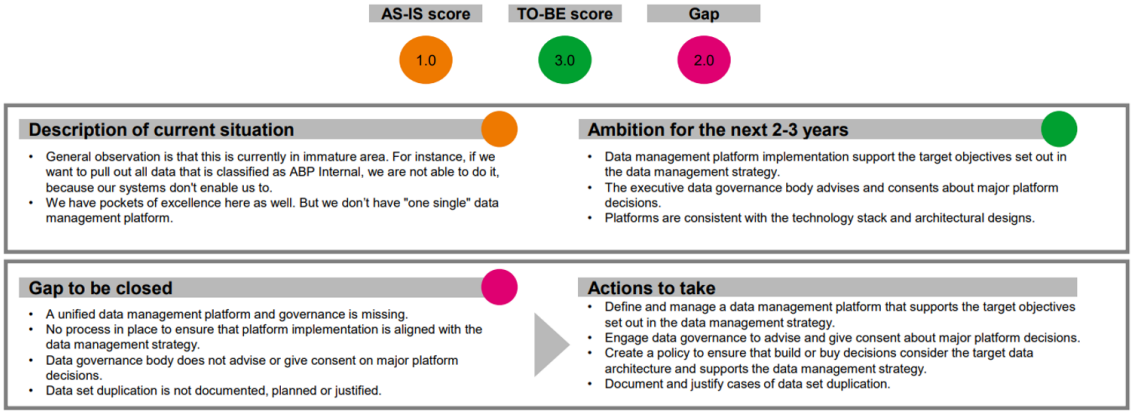


Figure 5.1: Data Management Platform Gap Analysis from the Data Foundation Blueprint [4].

The Data Factory within Aker BP is currently working on the actions above to close the gaps identified and establish a common data management platform. The DataOps Maturity Model described in Section 3.8 and the results from the GAP analysis in Figure 5.1 are closely interlinked. Based on the assessments and workshops conducted together with them for the analysis of this thesis it is clear that there is not one correct answer for data management tooling in Aker BP. All scenarios have advantages and disadvantages that need to be considered and weighted against each other. After evaluating a multitude of software solutions, as shown in 5.2, on the topics of data and metadata management, data quality and master data management for the selected areas of functional score, industry fit, technical fit and market reach. A shortlist was established and going forward the Data Factory are

evaluating the software solutions of Informatica, Collibra, Talend, Semarchy and Ataccama.

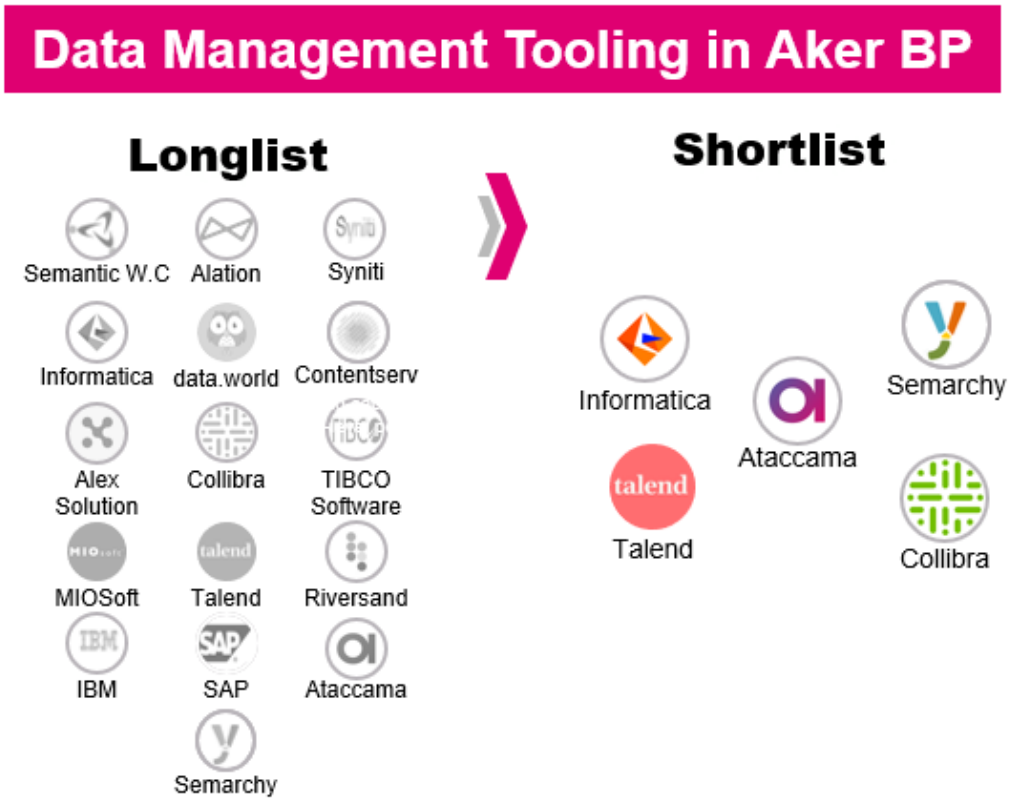


Figure 5.2: Data Management Tooling selection by Data Factory

The following sections will provide two examples and a brief overview of how the data management software solutions capabilities on the shortlist align with the DataOps vision of Aker BP and how they compare to DataOps platforms or tools already available in the market.

5.2 Data Management Tools

5.2.1 Collibra

Collibra helps organizations understand and manage their data assets by providing a data catalog platform and tool to create an inventory of data assets, capture metadata, and govern the assets. The main purpose of the tool is to help stakeholders understand what data assets exist, what they are made of, how they are being used, and if they are in regulatory compliance [17]. The four major functional areas of Collibra are:

- Data Catalog - To provide an inventory of data assets or products and allow users to find and access information quickly and easily. Users can search across several different aspect of the data assets.
- Data governance – Data governance in Collibra help create a common understanding of and sharing information about data assets such as technical metadata and additional user-added information.
- Data lineage – To see how data assets are created, transfered and transformed from system to system. Data lineage help data owners track where a data product comes from and how it is shaped.
- Data privacy – Allows security teams to create, manage and run policies to ensure data privacy and compliance.

5.2.2 Informatica

Informatica offer data integration products such data quality, master data management and data lineage. It focuses on extracting, transforming and loading data into the Informatica tool and for use cases such as:

- Migrating from legacy systems to a new database system.
- To move data from production systems to a data warehouse.
- Integrating multiple databases and file-based systems in one platform.
- As a data cleansing tool during transformation and loading of data.

Informatica also have the the necessary functions of a catalog to create a inventory and organize data assets, enriching metadata to your data and turning data into insights with built-in AI/ML applications. Also, the data catalog allows for tracking of data movement from to system to column level for detailed analysis and data lineage [31].

5.3 DataOps Tools

5.3.1 Cognite Data Fusion

Cognite Data Fusion (CDF) is Cognite's Software as a Service (SaaS) industrial DataOps platform. CDF stream data into the data model where data is normalized and enriched by connecting data sources of different types. To analyse and contextualize data in CDF, data integration pipelines between the existing data infrastructure and the data platform need to be established [8]. Contextualization tools make data more accessible and enables domain experts to map data from different sources to each other in the data model. The main functionalities of the DataOps platform can be summarized as:

- **Integrate Data** Prior to analysis and contextualization of data, efficient data integration pipelines between existing data infrastructure and CDF need to be present.
- **Contextualize Data** - Contextualization key to making data more accessible and making data do more.
- **Explore and Build** - Find, validate, and learn about the data available to build new solutions. Browse through available data sets and drill into details and related sources, for example, assets and time series with live data, interactive documents and 3D models.
- **Manage and Configure** - To control access to data and features in CDF, the organization define what capabilities users or application have to work with different data.
- **Trust and Security** - Cognite recognize the importance of data correctness

and preserving the original data and maintaining customer data security and privacy is critical.

5.3.2 DataKitchen

The DataKitchen DataOps Platform automates the key function of the development and production workflows to enable cross-functional teams to collaborate and deliver error-free, on-demand insight that lead to successful business decisions [15]. The platform is system agnostic and can be integrated with any analytic tool. The platform focuses on automated testing and monitoring and aim to improve data quality by catching data errors. In addition, the DataKitchen Platform fosters collaboration by providing a common place to work with data and an end-to-end view of the analytic process. It aims to increase visibility of the state of the data operations and provide metrics to show how teams are increasing collaboration, reducing errors, and speeding deployment times of new analytical insight [15].

5.4 Summary

As the gap analysis in Figure 5.1 highlight, Aker BP need to define and manage a data management platform that supports the target objectives of the DataOps function. Both data management tools of Informatica and Collibra have similar functionalities, and seem a good fit for Aker BP's requirements in a data catalog. Further analysis and testing should be initiated to land on the optimal solution for a data management tool and how it can be combined with the functionalities of a DataOps tool. As such, a DataOps tool could also be utilized as there are overlapping functionalities. Aker BP are already employing Cognite's CDF for data contextualization and liberation, and Cognite create data pipelines in order to enable this today. Potentially, Aker BP could utilize, or together with Cognite develop, further DataOps capabilities in CDF for it to become a fully-fitted DataOps tool. A few of the identified concepts in data management and DataOps platforms that will be directly connected to the to-be state of the DataOps maturity model in Figure 6.1 are:

- Building, testing, monitoring, and deploying data pipelines.

- Providing a flexible data catalog.
- Tracking and visualizing data lineage.
- Making data available for the business users.

Chapter 6

Development of DataOps Concept for Aker BP

This chapter aim to highlight the DataOps principles and concepts that will be imperative to Aker BP's vision of Industrializing DataOps and becoming a data-driven company. In addition, a high-level overview of the necessary efforts to increase D&W's DataOps maturity will be proposed together with an implementation plan in the following chapter. Section 3.8

In the analysis of Aker BP's DataOps maturity in Chapter 4, Aker BP was defined to be in the developing, level 2, stage of the model. At this stage some processes are in place to protect quality, data insights are used to inform business decisions and data is established to be an important asset in the company's strategy. Aker BP are taking strategic decisions that emphasize their trust in data being a key component to the company's success such as the data factory, CDF, OSDU and establishing a DataOps function to support these initiatives. To further develop the DataOps function, this research paper suggests to lift the process areas in Figure 6.1 to the next stage in the model and increase their maturity.

DataOps Maturity Model

Stage	1. Initial	2. Develop	3. Define	4. Measure	5. Optimize
Data Management & Analytics	• Ad-hoc ways of working. • Few datasets or products have been onboarded to the organization. • Data is used mainly for reporting purposes.	• Data insights are used to inform business decisions. • The organization can combine data from different sources. • Some established ways of working. • Data to be defined as an important asset.	• Dedicated teams focused on analytical value delivery and improvement. • DataOps practices are characterized for the organization and proactive. • Defined data catalogs used to manage data inventory and products.	• Data delivery and insight is measured and controlled. • Analytical insight and performance is measured and continuously improved. • Certain data management practices are automated across the organization.	• DataOps processes are fully automated, regularly improved and optimized. • New technology and analytics is developed upon an architecture designed for efficiency and large volumes of data.
Culture & Collaboration	• The use of data is based on individual efforts. • The business lacks a coherent data architecture and strategy. • No business unit is working on improving how the organization works with data.	• Business goals are known and communicated. • Activities are deliberate and documented. • Data is part of measuring results.	• DataOps practices are aligned with an enterprise-wide strategy. • Business users have cleaned data available to enable data-driven insight. • Deliberate ways of collaborating across business units.	• Decision makers are enabled with the results of data analysis to maximize business outcomes. • Service catalog of agreed technologies and high-level framework for teams to adopt new technologies.	• The business is constantly looking to leverage and combine data from new sources. • Automated AI/ML algorithms and data pipelines continuously improve business objectives.
Data Governance	• Governance is largely manual and lacks consistency.	• Processes are in place to protect data quality across the organization.	• Best practices are defined and evangelized, and teams enabled to evolve practices.	• There is universal confidence in the data and resulting insights.	• Data Governance is integrated into all business processes.

Figure 6.1: DataOps Maturity Model - Increasing Aker BP's DataOps maturity.

Reaching the 3rd stage in model above, Defined, require Aker BP to solve some the challenges they have related to all three process areas of data management and analytics, culture and collaboration and data governance. In the Defined stage Aker BP would have dedicated teams focused on analytical value delivery and improvement of both business and data management processes. The DataOps practices such as treating data as a product, close collaboration between data managers and consumers and established ways of providing analytical insight is established. Data catalogs or similar functionality is in place to manage the data inventory and products, and these are clearly communicated to the organization to ensure ease of access to data when needed.

For culture and collaboration, DataOps principles such as daily interactions with the customer and improving cycle times between analytic ideas to value are aligned with an enterprise-wide strategy to create a pull from the rest of the organization to utilize the DataOps function. Business users are continuously satisfied and their needs are met with having clean data readily available for either self-service analytics or finding new ways to collaborate on data with data disciplines. This collaboration across the business units is deliberate and defined so that each function working with data has a clear mandate and know what to expect from each other.

Finally, through data governance best practices, processes and standards are defined and evangelized. They are aligned with all of IT, DataOps and business requirements and are followed to increase the overall quality of their data products. This enables teams to continually evolve new practices and improve current processes as they are aligned with their ways of working.

The following sections will go into more detail of what Aker BP need to do to reach this maturity stage and some of the steps required to implement a successful DataOps practice.

6.1 Data & Analytics

Maximizing the usage of data through data platforms such as CDF or OSDU and building new technology on the foundation of these platforms sparked a lot of excitement in Aker BP. However, after the initial excitement it became clear to many that liberating large volumes of data from legacy systems into a common platform, and turning that data into insights and realizing value from it is a difficult task.

The analysis and discussions show that these problems are often related to poor data quality, absence of proper tooling, lack of collaboration, lack of standards and a misalignment between business and data disciplines. These problems combined have resulted in Aker BP not being able to extract the full potential of their data platforms and this section will explore some of the capabilities DataOps address in order to drive value from data more quickly and efficiently.

Figure 6.2 aim to show how the DataOps function in Aker BP will need to encapsulate all stages of working from data to supporting their input in existing master sources, liberating the data through APIs (or potentially OSDU and their standards), the data factory within DataOps is then responsible for managing the data catalogue, monitoring of the data and creating the data pipelines. The end product from the combined efforts of the data factory together with DataOps teams are then governed data products that can be used and consumed by business users, application and digital initiatives in EurekaX or I&P teams in the business.

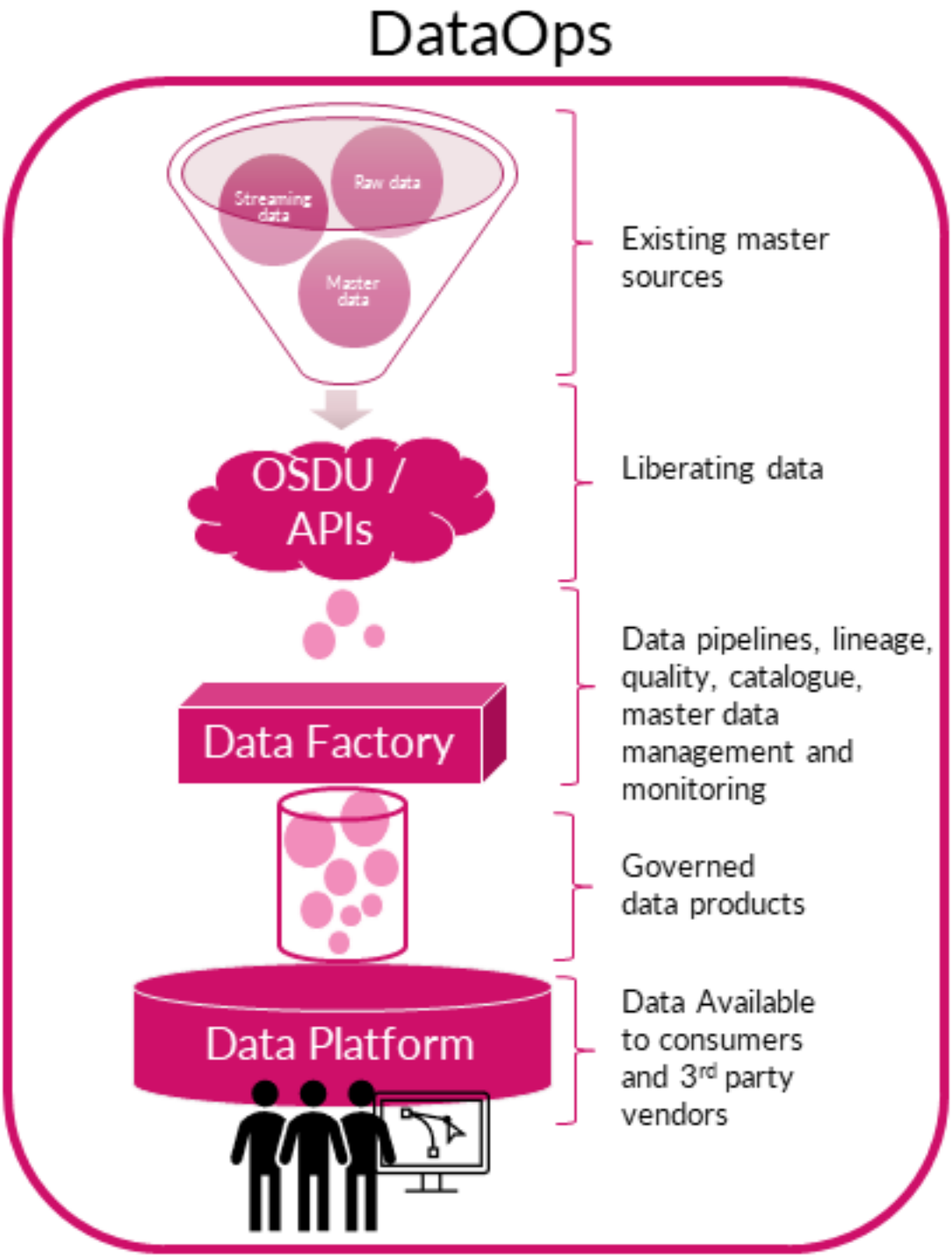


Figure 6.2: DataOps is key from liberation to consumption of data.

6.1.1 Data Monitoring

Data monitoring is the first practice in which data is checked against quality control rules to ensure its completeness, accuracy and validity. In DataOps the key behind data monitoring is observing data profiles over time and catching potential anomalies. It can be used to collect various metrics such as number of processes records, ranges of values for numeric columns, size of data and number of null or empty values. Aker BP can apply data monitoring at multiple stages of the processes described in Figure 6.2. Today, some of the data monitoring checks are done at the data platform level in Cognite's data pipelines for CDF. Going forward, data monitoring should be engraved in the DataOps work processes and combined with data lineage to monitor how data moves and transforms throughout the organization and between each step in the figure above.

6.1.2 Data Quality

Data monitoring in Aker BP can enable data managers and different data disciplines to have more insight about details of the data. However, monitoring data also requires rules and standards to be in place to define what is good quality. Aker BP need to establish business rules to check if data complies with pre-defined requirements, and firstly establish the pre-defined requirements. Most applications in the D&W business do not have pre-defined requirements, Aker BP specific mandatory attributes and governing documents. This gap need to be closed in order to monitor data quality and implement anomaly detection to set certain thresholds for the data.

The tool developed with Aker BP, Dr.Q is already being used in the organization today to monitor real-time WITSML data and its further use-cases should be investigated before acquiring a data quality monitoring tool.

6.1.3 Continuous Delivery

Tooling such as git should be applied to the data pipelines and PowerBI data flows already available from the data management team due to the increase in usage across the businett unit. Data testing and orchestration should be utilized for safe releases and testing before applying changes to production. The overall goal is to decrease the

time to implement and deploy new data pipelines and improve their overall quality and stability. A git-based system will allow for increased collaboration within the data disciplines and potentially reduce time-to-market for new data types over time.

6.2 Culture & Collaboration

Aker BP are heavily invested in developing new technology and digital advancements through EurekaX, the Data Factory was initiated based on the data foundation blueprint [4] to enable Aker BP to treat data as an asset, Improvement & Performance (I&P) teams require data easily available, the business is more invested in analytics than ever and finally a DataOps function has been established to transform data and information management teams since the start of 2022. As data has become so important to Aker BP within short time, there are an abounding amount of initiatives where DataOps need to be involved in some form or manner.

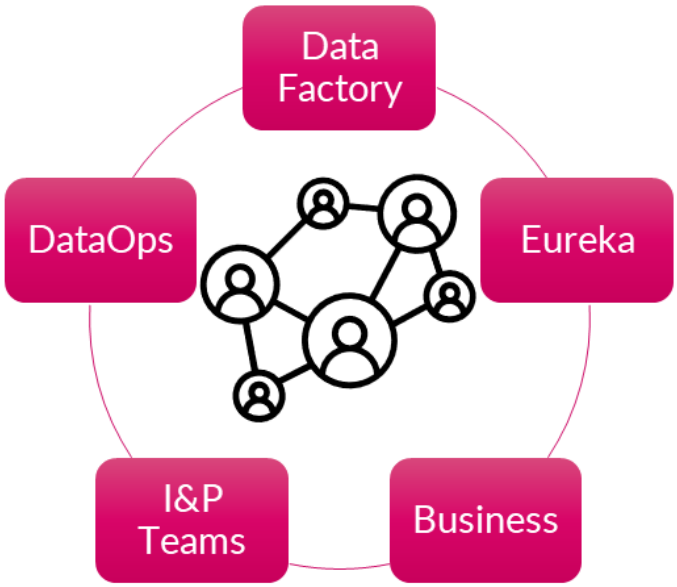


Figure 6.3: The five main functions involved in the DataOps concept presented.

The five functions described above, and shown in Figure 6.3, have a complex rela-

tionship and it will be important that the mandates of each function are clear to all. Below is a suggestion for a high-level summary of each function's main role and responsibility in a data-driven Aker BP.

DataOps Liberating consistent, standardized and quality data to all consumers.

Data Factory Establishing data products and automating data pipelines.

Improvement & Performance (I&P) Support the business through learning and data-driven improvement and providing data visualization platforms for effective decision making.

EurekaX The framework for all digital development initiatives and customer of new data products.

Business Show ownership towards data and look for opportunities to utilize and combine new data sources.

Each function play an important role in DataOps concepts presented and the collaboration between these function are critical to its success. One of the challenges is the interaction between DataOps and each of these complex interfaces. Figure 6.4 emphasize the difference between centralized teams supporting the company and distributed teams sitting close to the business itself. To take an example, today the D&W data management team is centralized in the DBT business unit, but located locally with the D&W business unit to be close to the customer and engineers. The newly established DataOps organization need to consider this when creating the organizational structure. The research and discussions carried out in this thesis suggest that being close to the business and understanding the business domain is critical for data management professionals. It also emphasizes how there could be a benefit in a hybrid solution where some teams, such as the D&W data management team are locally distributed, but there are data science specialists and a data governance function assisting with highly technical tasks or advising on best practices, respectively.

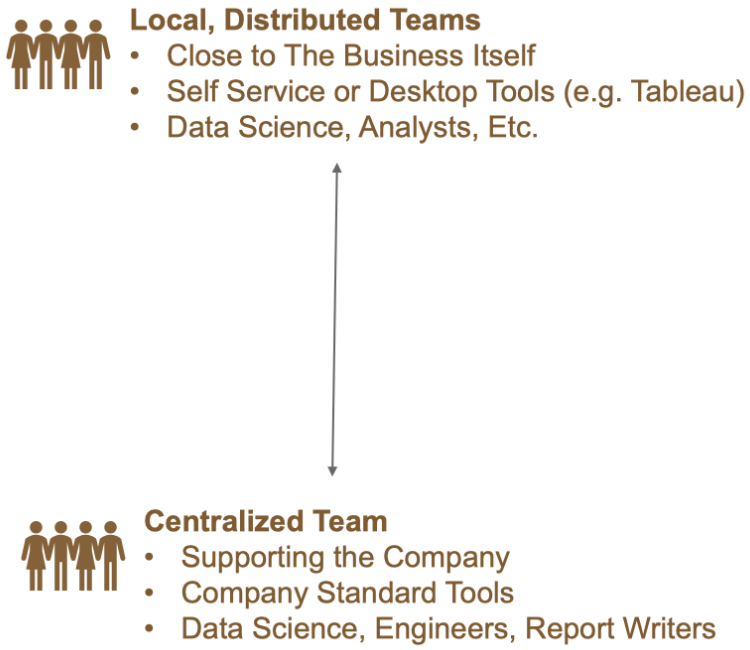


Figure 6.4: Centralized vs Distributed analytic teams [14].

6.2.1 DataOps Engineer

The DataOps engineer, as described in section 3.7, can be utilized in order to modernize and automate data management practices. As data managers struggle with overcoming the daily operational tasks of data management, there is little time for improvement of existing data management processes as seen in Figure 3.4. A DataOps engineer could be used to explore possibilities to replace manual procedures for data operations, establish a close relationship to the business and lift the capabilities of the data management teams.

Another potential improvement will be to implement DataOps teams. A DataOps team has all the capabilities of data managers, analytics, scientists, engineers in addition to a DataOps engineer as seen in Figure 6.5. The data engineer lays the groundwork of moving data from their respective applications into a data platform, a data warehouse or available through APIs or data flows. The analyst takes the data created by the engineer and provides analytics to business users and the data

scientist perform research, create algorithms and develop new models to solve problems. Finally, by adding the DataOps engineer the team is able to orchestrate and automate the analytics or data pipelines and break down barriers between operations and data analytics, improve work flows and processes and ensure a high level of quality of the data.

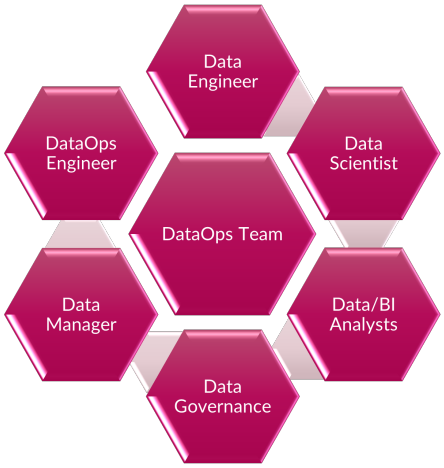


Figure 6.5: DataOps Team.

6.3 Data Governance

Since the start of this research a Data Governance Office has been established within DataOps and some of the focus areas will be to: Define and operationalize the data management strategy across the organization, establish governance roles, responsibilities and accountabilities, establish best practices, policies, procedures and standards within domain and monitor performance of governance in the data management function. In order for the Data Governance Office (DGO) to be successful, it will be important to tackle the various interfaces in Figure 6.3.

Two aspects of data governance that are critical to Aker BP's DataOps concept are the data catalog and data lineage. Both enable data scientists, analysts, and engineers to quickly find required datasets and how they have transformed or traveled. Tools like Collibra and Informatica are potential starting points in implementing this capability. Figure 6.6 show how a data governance platform encapture data

catalogs, data lineage, data glossaries and dataset profiles.

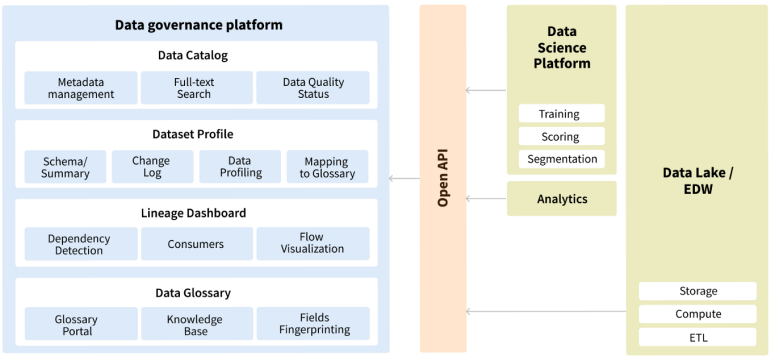


Figure 6.6: A Data Governance Platform with the capabilities of a data catalog, glossary, lineage and dataset profiles [21].

6.4 Tools

Based on the DataOps maturity level in Figure 6.1 Aker BP are not mature enough for fully automated data pipelines across the organization. Rather, the focus should be on maturing to the next level by implementing certain functionalities of a DataOps tool such as providing a flexible data catalog and making data available for the business users. In addition, Aker BP should look to find use cases for building, testing, monitoring and deploying data pipelines and tracking data lineage before further adoption across the organization. A common data management or DataOps tool with the relevant features of a data catalog and data lineage such as Collibra or Informatica should be considered. This will allow data managers to more easily monitor data quality, and business users will be able to navigate to the required data types they require with ease. In addition, as data is made available through a catalog Aker BP should also include the data source and quality which will highlight the issues related to undefined master data sources and poor data quality in key systems.

6.5 Summary

As discussed, DataOps as a data management practice focuses on improving communication, integration and automation of data flows between data managers and data consumers across an organization with the goal of enabling data-driven decisions through having data readily available with good quality. Aker BP have already embarked on their DataOps journey and DataOps is one of the key functions in the Digital Business Transformation (DBT) business unit. However, DataOps practices are still in its infancy and both the data foundation blueprint [4] and this research paper highlight the gaps and issues Aker BP have related to data management and data governance across the organization. It is clear that Aker BP need to shift from application focused data management to managing data as an asset to successfully deliver on their data-driven ambitions. Master data management, data lineage and data catalogs are all important aspect on this journey, and the current research show that DataOps as a framework for these concepts has a lot of potential if executed correctly.

Chapter 7

Implementation Plan & Critical Success Factors for Aker BP

7.1 Implementation Plan - Maturity level 3

When embarking on the DataOps journey, Aker BP intend to succeed in improving their ability to make data-driven decisions and improve business processes. Their goal is to reach the 5th maturity level in the DataOps Maturity Model in Figure 3.5. However, the final maturity level can only be achieved by completing each maturity level step-wise, and firstly Aker BP need to reach DataOps maturity level 3. An implementation plan has been developed through this research and Figure 7.1 identify certain key areas of improvement to increase the maturity level from level 2 to 3 in Aker BP.

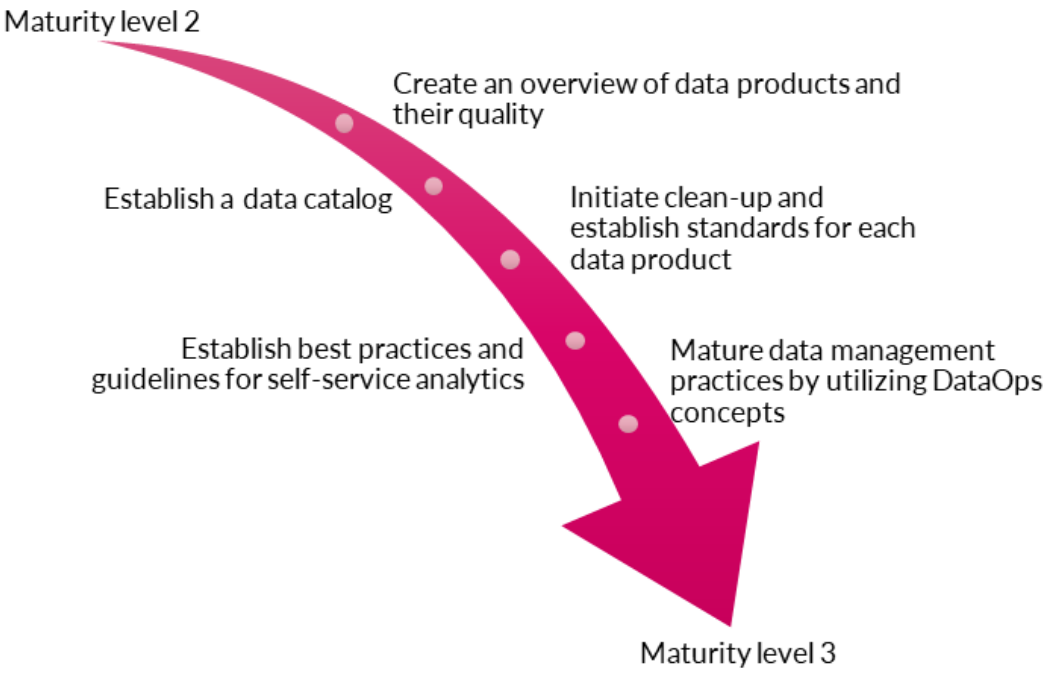


Figure 7.1: DataOps Maturity Model - Where Aker BP can increase their DataOps maturity.

Creating an overview of data products and their quality will be crucial in order to drive transparency in the data available in the organization or business unit. The primary objective of DataOps is to increase the quality and usage of data in the business, and as such communication of which data types or products are available in what systems, with what quality, how it can be extracted and their respective owner will be essential. This is the first step necessary to create an enterprise data catalog and does not require a data management platform, but could be in the form of a Power BI dashboard or information page for transparency.

Establishing a data catalog can be performed when the overview is available. The data catalog should provide an inventory of data products that allow users to find and discover the right data to use for their purpose. The data catalog is more comprehensive than the overview and can include capabilities to automate the delivery of relevant data and govern the use of data in compliance with the

organizational requirements.

Initiate clean up and establish standards for each data product. Aker BP's ability to create value from digital projects and liberation of data to business users is no better than the quality of the data being made available to the systems and users. In order to close the gap of poor data quality residing in various applications, Aker BP need to establish data standards and requirements for each products. Only then can data be collected with the right quality, and certain systems will need to have clean up projects initiated to ensure that the data in the system is to be trusted.

Establishing best practices and guidelines for self-service analytics is essential for self-service users and analytic teams in the organization to be effective. Different analytic or modelling tools require best practices on how they shall be used, what tool to be used for what, who to contact for assistance and how the results created can be transferred from development to being operationalized. This improvement area will enable clear ways of working with data and analytics and improve data governance in the organization.

Mature data management practices by utilizing DataOps concepts. A key area of improvement highlighted in this thesis is how Aker BP perform data management today and experience data errors. Establishing data products, automating delivery and quality assurance of data, and improving collaboration with the business customer are all key components of maturing DataOps and data management within Aker BP.

7.2 Critical Success Factors

Aker BP show a willingness to experiment with new technologies and concepts through EurekaX, CDF, Data Factory and already undergoing an organizational restructure to go from data and information management to Data Operations. Ways of working, interfaces to the other relevant functions, and tools and technologies are still being defined and certain concepts have been proposed in this thesis. This section will cover some critical success factors related to the technical, operational, and organizational factors Aker BP should consider on their DataOps journey.

7.2.1 Technical Factors

Data Products

Aker BP are not mature in treating data as a product, and the same data product may have multiple versions of the truth living in different systems and projects. When starting to define and develop data products it will be critical to ensure that these products are reuseable across the organization. Every data product needs an owner and Aker BP need to equip these owners with the skills required to take full responsibility for the product in the role of data product owner or manager.

Technical adaption

DataOps introduces new concepts such as data pipelines, CI/CD, data management tools, data lineage and data mesh to data management and analytics. It will be critical to consider the change management of new work processes, automated business processes and the technology introduced. Data managers and analytic teams need to be informed and interested in the use of the new concepts and data products to ensure adoption across the organization.

Knowledge transfer

The data management role today is based on individual efforts and experiences. Aker BP will need a good system in place to share knowledge across teams and within the teams, document their efforts and communicate the deliverables of data management teams.

Tool selection

Efforts are still needed in order to define an appropriate data management or DataOps platform for the organization. It will be important to do a thorough analysis of the necessary functionalities needed, how they will be used and how the DataOps organization will adopt and integrate a new tool into the existing work processes.

7.2.2 Operational Factors

Defining goals

Clear goals need to be defined for the various DataOps functions and KPIs should be created to measure deliverables and understand the performance and health of the business unit. This will be important in order to make adjustments to the function's execution of their strategic goals and adopt to changing priorities in the organization.

A clear mandate

With increasing organizational complexity and multiple parts of the organization working on seemingly similar functionalities and concepts, it will be more important than ever for each function to have its mandate. This can improve the collaboration between each function and will strengthen the understanding of each interface when working with data.

Ability to operationalize and scale data products

Aker BP are still in the infancy of treating data as a product instead of managing the data in their respective applications. It will be critical that Aker BP and their DataOps function are able to communicate what a data product is, what the vision is and why they are looking to mature this area. This requires buy-in from the data managers in the organization today and close collaboration with the data factory and the data products defined in their projects will be crucial.

7.2.3 Organizational Factors

Organizational adaption

Digital transformation is about more than just technology and requires a culture capable of supporting it. The participants involved in this study indicate that there has been a large cultural change since the origin of Aker BP and after initiating EurekaX as work processes are becoming more digital and people see the need for new technologies and increased data quality. If Aker BP are to be successful with their digital and data initiatives it will be critical that the culture in the organization

is aligned with the companies strategic goals and digital transformation.

Collaboration between business units

As DataOps is a new function entering a complex organizational environment it will be important to define in what areas DataOps should collaborate with other business units. In what scenarios should the function provide distributed teams assigned to projects or business support, and what areas do the other business units expect to have centralized centers of excellence to provide standards and guidelines. Each business unit today have different requirements to the DataOps organization and it will be important to prioritize and standardize across the organization as Aker BP continuous to grow.

Onboarding of new resources

The increased focus on data and the introduction of DataOps and Data Governance in Aker BP has led to onboarding of new resources, and the acquisition of Lundin will lead to further growth in the organization over a short period of time. Aker BP need to clearly communicate the vision and business value of having a DataOps function in the organization and encourage new hires to participate in continuous improvement of the function. In addition, Aker BP have many projects progressing in parallel and understanding the interfaces between each is complex. Clear communication and information easily available will be key to onboarding the new hires in Aker BP.

Chapter 8

Conclusion

The purpose of this thesis was to investigate one of the key frameworks Aker BP aim to utilize when working with data and analytics to become a data-driven company, DataOps. The main problem to investigate was how DataOps as a framework and concept can support Aker BP's strategic priority to become a data-driven company. The outcome and primary motivation of the thesis work was to provide additional support and authority on DataOps approaches and practices to Aker BP in their efforts to mature their data capabilities, culture and processes.

In Aker BP today, data management efforts are limited and based on ad-hoc need and fast changing priorities. Execution is highly manual as sufficient structure and tool support is missing. The pressure on the data management function to verify and clean up data sets is becoming larger each year as more digitalization projects are realised through EurekaX, Aker BP's digital lab.

As Aker BP aim to build digital capabilities and implement new ways of working where decisions are driven by data. In Aker BP today, data management efforts are limited and based on ad-hoc needs and changing priorities. The execution of data management tasks are highly manual and a sufficient framework and tool support is missing. As such, maturing their data management and data governance capabilities is of high priority and this research highlight some of the issues Aker BP face when working with data today, in addition to the steps that need to be taken to improve

these functions.

DataOps has the potential of accelerating Aker BP's data transformation journey by creating a culture of federated data responsibility, maturing data management capabilities and building a common data management and governance model across the organization. It will enable lead Aker BP towards a data-driven company with efficient and reliable data operations, with a continuous pursuit of improving data flows and data product deliveries on an integrated set of data platforms. Succeeding on the DataOps journey Aker BP have began will remove data silos reducing the availability and accessibility of data and improve the trust in data, further enabling remote work, predictive maintenance and unattended offshore facilities. DataOps is as much about people as it is about tools and processes and focuses on improving the communication and integration of data flows between data managers and data consumers across an organization.

The DataOps maturity model developed in this research has Aker BP classed as Stage 2 'Develop' where Stage 5 'Optimize' is the company's long-term ambition. The thesis focuses on the necessary efforts to increase Aker BP's maturity to Stage 3 'Define' and proposes a clear implementation plan for the necessary efforts to reach this stage, such as establishing data products and establishing their standards, organizing these in a data catalog and enabling self-service analytics in the business with best practices and guidelines for working with data. The DataOps maturity model is generic and can be used to classify a company's readiness to apply DataOps concepts and identify potential improvement areas.

8.1 Learning points

By investigating and analysing how Aker BP perform data management and data governance in the organization today it is clear that there is a gap to be closed to reach Aker BP's ambitions. DataOps' principles and purpose align well with the participants impression of how an organization should treat its data and whilst it is an emerging concept it is viewed as solution to many of their problems. Establishing data products, treating data as an asset, and improving collaboration with the business customer are seen as key components in succeeding with DataOps.

8.2 Challenges

Data managers today struggle with being occupied in ad-hoc data deliveries or cleaning and operational day-to-day tasks. As such there has been little time to improve data management tasks. The implementation of DataOps concepts require Aker BP to lift their data management function, and for the team to drive the change towards data products, improved collaboration and increase availability of data prioritization have to change or the function has to grow. The following challenge to this issue is how data management in the E&P sector is very specific to its applications, and domain knowledge in addition to knowledge of how its application is used is critical. This experience can take years to acquire, and Aker BP may need to develop cross-functional data management professionals that understand both a business domain and the technical aspects of the data management. In addition, there are technical challenges related to key systems in the organization not being able to liberate its data through open APIs or similar due to their deficiencies.

8.3 Further Research

DataOps is an emerging concept that is likely to continue to evolve in the years ahead. This thesis focused on the principle that DataOps' goal is to liberate data from its sources to its consumers, and the activities of data pipelines and automating business processes were mentioned but not thoroughly analyzed. Researching this aspect of DataOps further, analyzing the DataOps tools and their function in practice and experimenting with data pipelines will be an important next stage. A data analytics project where a DataOps methodology and tools are used in practice could be a starting point to prove the benefits of adopting DataOps practices when working with analytics throughout the organization.

In addition, the implementation plan need to be acted upon and can be used as a basis for future work to implement certain aspects of DataOps. A compatibility rating based on the functional fit, security requirements and quality evaluations of the various DataOps tools would help DataOps practitioners make informed decisions.

Assigning data ownership has been emphasized in this thesis and an analysis of

how clear ownership within the organization will enable ongoing management and changes to data to meet business needs would be beneficial. Furthermore, establishing a review process to make sure data governance, standards, policies and processes are being followed and a process to establish these should be in place. Finally, improving data quality by enforcing and complying to the defined processes, policies and standards that set clear data quality requirements need to be executed.

References

[1] Accenture. *Aker BP and Accenture Team Up on Data Factory in the Cloud to Improve Oil and Gas Operations*. URL: https://newsroom.accenture.com/news/aker-bp-and-accenture-team-up-on-data-factory-in-the-cloud-to-improve-oil-and-gas-operations.htm.

[2] Aker BP ASA. *Aker BP acquires Lundin Energy's oil and gas business*. URL: https://akerbp.com/borsmelding/aker-bp-acquires-lundin-energys-oil-and-gas-business/.

[3] Aker BP ASA. *Our approach to sustainability*. URL: https://akerbp.com/en/sustainability/.

[4] Aker BP. "Data Foundation Blueprint". Aker BP. 2021.

[5] Core BTS. *Power BI Governance Delivery Strateg Licensing*. URL: https://corebts.com/blog/power-bi-governance-delivery-strategy-licensing/.

[6] Eran Strod Chris Bergh and James Royster. *Recipes for DataOps Success*. DataKitchen, 2021. URL: https://datakitchen.io/recipes-for-dataops-success-guide-to-dataops-transformation/.

[7] CIO. *Data governance - A best practices framework for managing data assets*. URL: https://www.cio.com/article/202183/what-is-data-governance-a-best-practices-framework-for-managing-data-assets.html.

[8] Cognite. *Cognite Data Fusion*. URL: https://docs.cognite.com/cdf/.

[9] Cognizant. *Data Platform*. URL: https://www.cognizant.com/us/en/glossary/data-platform.

[10] Collibra. *What is Data Lineage*. URL: https://www.collibra.com/us/en/blog/what-is-data-lineage.

[11] The Journal of Commerce online. *Digitalization: Understanding the Role of Data Management*. URL: https://www.joc.com/digitalization-understanding-role-data-management.

[12] the Data Governance Institue. *Goals and Principles for Data Governance*. URL: https://datagovernance.com/the-data-governance-basics/goals-and-principles-for-data-governance/.

[13] Den Norske Dataforening. *How Data Lineage enhances & accelerates all Data Governance initiatives*. URL: https://www.dataforeningen.no/how-data-lineage-enhances-accelerates-all-data-governance-initiatives/.

[14] DataKitchen. *DataOps is NOT just DevOps for Data*. URL: https://medium.com/data-ops/dataops-is-not-just-devops-for-data-6e03083157b7.

[15] DataKitchen. *DataOps Platform*. URL: https://datakitchen.io/platform/.

[16] DataKitchen. *What is a DataOps Engineer*. URL: https://datakitchen.io/what-is-a-dataops-engineer/.

[17] Datameer. *What is Collibra*. URL: https://www.datameer.com/what-is-collibra/.

[18] Dataversity. *Understanding Data Management Tools*. URL: https://www.dataversity.net/understanding-data-management-tools/.

[19] Zhamak Dehghani. *Data Mesh*. URL: https://www.thoughtworks.com/radar/techniques/data-mesh.

[20] DNV. *Normally Unattended Installations*. URL: https://www.dnv.com/article/normally-unattended-installations-208008.

[21] Grid Dynamics. *5 Steps to Implementing a Successful DataOps Practice*. URL: https://medium.com/codex/5-steps-to-implementing-a-successful-dataops-practice-f553effd7367.

[22] Ericsson. *How digital twins in the oil and gas industry can modernize your business*. URL: https://www.ericsson.com/en/blog/2021/11/how-digital-twins-in-the-oil-and-gas-industry-can-modernize-your-business.

[23] Gartner. *DataOps*. URL: https://www.gartner.com/en/information-technology/glossary/dataops.

[24] Gartner. *Digital Twin*. URL: https://www.gartner.com/en/information-technology/glossary/digital-twin.

[25] Gartner. *Gartner Hype Cycle for Data Management Positions Three Technologies in the Innovation Trigger Phase in 2018*. URL: https://www.gartner.com/en/newsroom/press-releases/2018-09-11-gartner-hype-cycle-for-data-management-positions-three-technologies-in-the-innovation-trigger-phase-in-2018.

[26] Gartner. *Master Data Management (MDM)*. URL: https://www.gartner.com/en/information-technology/glossary/master-data-management-mdm.

[27] Gentackle. *The Impact of Industry 4.0 and Digital Transformation on Asset Management: The Basics You Need to Know*. URL: https://www.gentackle.com/digital-transformation-asset-management/.

[28] GEP. *The increasing popularity of digital twins in oil and gas*. URL: https://www.gep.com/blog/mind/the-increasing-popularity-of-digital-twins-in-oil-and-gas.

[29] IBM Big Data & Analytics Hub. *3 reasons why DataOps is essential for big data success*. URL: http://www.ibmbigdatahub.com/blog/3-reasons-why-dataops-essential-big-data-success.

[30] IBM. *Data Catalog*. URL: https://www.ibm.com/topics/data-catalog.

[31] Informatica. *Data Catalog*. URL: https://www.informatica.com/products/data-catalog.html.

[32] Informatica. *What is Metadata Management*. URL: https://www.informatica.com/nl/resources/articles/what-is-metadata-management.html.

[33] BI-Insider. *Master Data Management*. URL: https://bi-insider.com/portfolio-item/master-data-management/.

[34] Dama International. *DAMA-DMBOK Data Management Body of Knowledge (2nd Edition*. Technics Publications, 2017. URL: https://www.dama.org/cpages/body-of-knowledge.

[35] ISO. *Benefits of Standards*. URL: https://www.iso.org/benefits-of-standards.html.

[36] ISO. *ISO 8000-61:2016(en)*. URL: https://www.iso.org/obp/ui/#iso:std:iso:8000:-61:ed-1:v1:en.

[37] ISO. *ISO/CD TR 15926-14*. URL: https://www.iso.org/standard/75949.html.

[38] ISO. *ISO/TS 15926-4:2019*. URL: https://www.iso.org/standard/73830.html.

[39] The DataOps Manifesto. *DataOps Principles*. URL: https://dataopsmanifesto.org/en/.

[40] Dominik T. Matt et al. "Industrial digitalization. A systematic literature review and research agenda". In: *European Management Journal* (2022). ISSN: 0263-2373. DOI: https://doi.org/10.1016/j.emj.2022.01.001. URL: https://www.sciencedirect.com/science/article/pii/S0263237322000019.

[41] McKinsey. *How companies can use DataOps to jump-start advanced analytics*. URL: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/tech-forward/how-companies-can-use-dataops-to-jump-start-advanced-analytics.

[42] The American Society of Mechanical Enginees. *5 Lean Principles Every Engineer Should Know*. URL: https://www.asme.org/topics-resources/content/5-lean-principles-every-should-know.

[43] Sandra Melo. *What are benefits of ISO*. URL: https://datascope.io/en/blog/what-are-benefits-of-iso/.

[44] Microsoft. *What is PowerBI*. URL: https://powerbi.microsoft.com/en-my/what-is-power-bi/.

[45] Standards Norway. *Z-Technical Information*. URL: https://www.standard.no/en/sectors/energi-og-klima/petroleum/norsok-standard-categories/z-technical-info/.

[46] Oracle. *What is Data Managmeent*. URL: https://www.oracle.com/uk/database/what-is-data-management/.

[47] OSDU. *The OSDU Forum will enable Operators to be more efficient*. URL: https://osduforum.org/about-us/communities/operators/.

[48] Norwegian Petroleum. *Fields*. URL: https://www.norskpetroleum.no/en/facts/field/.

[49] Data Quality Pro. *Beginners Guide to Master Data Management (MDM)*. URL: https://www.dataqualitypro.com/blog/beginners-guide-to-mdm-master-data-management.

[50] READI. *ISO 15926-14:2020(E)*. URL: https://readi-jip.org/wp-content/uploads/2020/10/ISO_15926-14_2020-09-READI-Deliverable.pdf.

[51] MIT Technology Review. *DataOps and the future of data management*. URL: https://mittrinsights.s3.amazonaws.com/dataops.pdf.

[52] Satori. *All you need to know about DataOps tools*. URL: https://satoricyber.com/dataops/all-you-need-to-know-about-dataops-tools/.

[53] BV Solutions. *How Digitalization is Winning Over the Maritime Industry*. URL: https://www.bvsolutions-m-o.com/magazine/how-digitalization-winning-over-maritime-industry.

[54] Frost & Sullivan. *Global Oil and Gas Automation Growth Opportunities*. URL: https://www.researchandmarkets.com/reports/5527058/global-oil-and-gas-automation-growth-opportunities?utm_source=CI&utm_medium=PressRelease&utm_code=j62zj7&utm_campaign=1655069+-+Global+Oil+and+Gas+Automation+Market+Report+2021%5C%3a+Growth+Opportunities+in+AI+as+a+Tool+to+Achieve+Sustainability%5C%2c+Robotics+for+the+Upstream+Sector%5C%2c+%5C%26+IoT+for+Boosting+Efficiency&utm_exec=chdo54prd.

[55] Svitla. *18 Best Data Management Tools*. URL: https://svitla.com/blog/18-best-data-management-tools.

[56] CMMI Product Team. *CMMI for Software Engineering, Version 1.1, Staged Representation (CMMI-SW, V1.1, Staged)*. Software Engineering Institute, Carnegie Mellon University, 2002. URL: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=6217.

[57] TechTarget. *Data Quality Definition*. URL: https://www.techtarget.com/searchdatamanagement/definition/data-quality.

[58] Terminusdb. *What is a Data Product*. URL: https://terminusdb.com/blog/what-is-a-data-product/.

[59] Cornell University. *Metadata and describing data*. URL: https://data.research.cornell.edu/content/writing-metadata.

[60] G. Vial. "The Journal of Strategic Information Systems". In: (2019), pp. 118–144.