AALBORG UNIVERSITY
COPENHAGEN

Semester: 4th Semester

Title: The Development of Data Strategy Framework: A Case Study

Aalborg University Copenhagen
A.C. Meyers Vænge 15
2450 København SV

Secretary: Maiken Keller

Project Period: 01-02-2018 – 07-09-2018

Semester Theme: Governance and Strategies

Supervisor(s):
Iwona Maria Windekilde

Project group no.: ICTE4BUS
4.12

Members :
Gintare Simkute
(do not write CPR.nr.):

Abstract:

This research discussed about the importance of data strategy and its alignment across the organization. The purpose of the research is to develop a data strategy framework that would be applicable for the case study. The research is based on the literature review and the empirical data from the interviews within the company of case study. During literature review, various existing strategy alignment theories and existing data strategy framework have been investigated and applied to the case study by the use of conceptual framework. Graphical representation of the data strategy framework is developed for the case study that shows the process of building data a comprehensive data strategy in the organization.

Key words: Data Strategy – Strategy Alignment – Conceptual Framework

Pages: 78
Finished: 07-09-2018

All group member are collectively responsible for the content of the project report. Furthermore, each group member is liable for that there is no plagiarism in the report. Remember to accept the statement of truth when uploading the project to Digital Exam

AALBORG UNIVERSITY

THE DEVELOPMENT OF DATA STRATEGY FRAMEWORK: A CASE STUDY

Master Thesis writen by

Gintare Simkute

2018, September

Alborg University, Copenhagen

Table of Contents

List of Figures

List of Tables

Acronyms

AI - Artificial Intelligence
API - Application Programming Interface
BD - Big Data
CARG - Compound Annual Growth Rate
CDO - Chief Data Officer
CIO - Chief Information Officer
CEO - Chief Executive Officer
DWH - Data Warehouse
EU - European Union
GDPR- General Data Protection Regulation
ICT - Information and Communication Technology
IRS - Internal Revenue Service
IS - Information Systems
IT - Information Technology
ML - Machine Learning
MVOT - Multiple Versions of The Truth
RMF - Risk Management Framework
SAM - Strategic Alignment Model
SSOT - Single Source of Truth
US - United States
SAVV BIGD - A Strategic Approach for Visualizing the Value of Big data
SVDS - Silicon Valley Data Science
SWOT - Strengths, Weaknesses, Opportunities, and Threats
EY - Ernst & Young
ROI - Return On Investment
AIG - American International Group
HL7 - Health Level Seven International
FIBO - The Financial Industry Business Ontology
NIEM - National Information Exchange Model
ANSI - American National Standards Institute
ISO - International Organization for Standardization
NIST - National Institute of Standards and Technology
IRS - Internal Revenue Service
TDWI - Transforming Data with Intelligence

VR - Virtual reality

NASA - National Aeronautics and Space Administration

SOTA - State of the Art

1. Introduction

Data has been the topic that everyone is becoming more and more interested in. Especially for the past few years, data caught more attention with the rise of the machine learning, data-driven decisions [1] and General Data Protection Regulation in EU. With the rise of data comes the challenges of managing it effectively, not only to comply with the standards but to add the value to the company [2]. Companies like Airbnb, Spotify and Amazon are using their data to make their services more efficient and generate even more revenue.

However, according to Gartner [3], only 9 percent of companies are able to use their data in the very effective way that would bring them value and benefits. “Cross-industry studies show that on average, less than half of an organization’s structured data is actively used in making decisions—and less than 1% of its unstructured data is analyzed or used at all. More than 70% of employees have access to data they should not, and 80% of analysts’ time is spent simply discovering and preparing data” [4]. Moreover, piles of valuable data are lying around not being used and even it comes as a burden for some companies as they face the data storage issues and expenses. By not managing and utilizing the data, organizations are losing their competitive advantage in the market. Accordingly, some actions should be taken into place.

In order to get the value from the data, data needs to be managed, The first step for companies to make their data usable and get any value from it is to develop a data strategy that would include organizational structure, infrastructure for data storage and data flows and data processing/business intelligence tools. Especially now, with the rise of new roles as chief data officer (CDO), the need for data strategy and data organization is becoming more and more important. Also to ensure the continuity of decisions and actions in the company as CDOs stay within the role just 2.4 years on average [4]. Data strategy and data management plan would ensure the consistency and coherent decisions toward data in the organization.

1.1 Background and Motivation

The fascination about data and the ways to use it has inspired the main subject of this Master Thesis, data and getting data to work for the company by creating guidelines on how to create a data strategy.

Why is data strategy so important to create in the first place? To begin with, data strategy should be created to help to manage the data across the organization, as it is impossible to keep data in sync across the company without having a strategy in place [5]. As the concept of the data strategy is relatively new, however, the big corporations already had data strategy

over than 10 years ago and were able to create sophisticated reporting because of it [6]. It was the first stage of the data strategy, also called as enterprise data strategy and was focused mostly on the reporting on the financial performance [6].
The data strategy view has changed dramatically over the past 10 years with the growth of data volumes and complexity, and the development of new, more sophisticated data tools and technologies. The purpose of using data also has a change over those 10 years, the companies now are using data in order to understand the needs and behavior of their customers, improve their everyday business operations or find new ways of having more revenue streams [6]. So the data strategy should also change and adapt to the current trends.

This Master Thesis includes the study case of Unity Technologies. It is a company that has been established 12 years ago and that has become a “Silicon Valley Unicorn”, meaning that it is a startup that is evaluated of $1 billion or even more” [7]. Unity Technologies is the company that sells subscriptions to video games creation engine, and that generates huge amounts of data every single day. The company gets data from the direct users and from the users of their customers. The need for an effective data strategy is needed for this company to grow even bigger in revenue and market share. The company was chosen for the case study due to the enormous amount of data that are being generated and due to the fact that the company is still at its growing phase and does not have clear traditions of creating strategies. The need for an effective data strategy is needed for this company to grow even bigger in revenue and market share. Well-crafted data strategy can increase data value for the company so that it could be monetized or effectively used in business operations. The data strategy for Unity Technologies could be applicable for the similar profile companies as well.

1.2 Problem Area

In this Master Thesis, the data part of strategy represents the big data concept, as the study case is a company that deals with high volumes of data every day. However, in the literature, the strategy is being referred to as just a data strategy, big data is only the characterization of the data that needs to be managed. This Master Thesis will follow the same naming as it is in most of the literature: data strategy. However, the characteristics of big data still need to be addressed.

1.2.1 Data and the Organization

What is data and why it is so important to the company? When we talk about data, we usually talk about big data, especially in the fast-growing companies. "Big data refers to the dynamic,

large and disparate volumes of data being created by people, tools and machines" [8], it is named as big data due to its huge volume and complexity. Usually, big data is characterized into 4 categories called as 4 Vs [8]:

- Volume: vast amount of data
- Variety: data is being generated from many different data sources and types
- Velocity: data is being generated constantly and at the very fast speed
- Veracity: refers to the quality and accountability of data, as it has a lot of noise and is being generated through different sources.

Majority of companies are generating big data every day. The data varies from company to company, they get data from the customers' account, customer behavior online, purchase orders, systems, and applications that run in the company. All of this data could be used to make business operations more efficient and make data-based decisions within the company and also to impact the customer with his decisions in the purchase [9].

However, only a small part of the companies are exploiting all the benefits of the data they can access to improve their business [3]. So why data is so important to the companies and why it needs to be used to its full potential? Data has the essential role to play in the improvement of the company's services and products, cost reduction, managing risks and increasing effectiveness in business processes. Data "flows through every business process - it shapes customer interactions, guides product design, and development, and facilitates strategic plans and decisions" [10]. Basically, by not managing data, companies would not be able to operate functionally in today's constantly changing market and the data that companies have could easily lose the value. Data has become so complex in the past years that it needs a separate strategic model to manage it. “Strategically, data serves as a critical component of the firm’s operational backbone for customer engagement and digitized solutions” [11], which makes data strategy a significant part for the organization.

1.2.2 Strategy

This Master Thesis is about data strategies. In order to analyze data strategies at first the basic concept of strategy needs to be understood, what the strategy is and why it is needed in the organization?

Strategies have been used mainly for the military purposes to develop a plan of how to destroy or defend against the enemy. This concept of the strategy has been adopted and widely used in the business world to conquer the market share and defeat the competitors [12]. The need for the strategy in the business world comes from the technological and societal changes that impacted the competition and demand in the market, so the strategy directs companies how

to achieve agreed upon goals and objectives [13]. A strategy is a framework that is combined with an interrelated set of actions that focus on long-term plans to strengthen the organization in relation to the competitors considering its strong sides and limitations [14].

Strategy in the business world is very commonly related to the three essential organizational strategic components: vision, mission and core values [15]. Those components explain what are the elements that define the business purpose, the business objectives, future vision and the ways to reach it. The sequence of planning the strategy is illustrated in Figure 1. [16]

Figure 1. The vision, mission, objectives, strategy and action pyramid [16]

- The vision of the organization describes a very high-level purpose of the existence of that organization. It shares the beliefs and values and the future vision of the company.
- The mission statement communicates on the high level what the organization is trying to achieve and why it is important. The mission statement is similar to the vision statement, it is just more action-oriented and more concrete [17].
- The objectives - this part describes clear goals of the company together with the success factor and the timeline to accomplish it [16]. The objectives are created to complete the mission statement.
- The strategy - this part describes how those objectives will be achieved. The strategy includes a variety of methods and procedures together with many different people, resources developing them [17].
- The action plan is the last part of the strategy pyramid, which details the actions that are needed to take in order to achieve each objective. It is the roadmap of the strategy that describes in great detail what needs to be implemented and how it needs to be implemented including the action steps, resources, timelines and constraints [16].

This strategic planning helps organizations to clear their vision, set the goals and objectives, construct the way how to achieve all that in a defined way [17].

1.2.3 Data strategy

“To effectively achieve business goals, organizations must understand the data they have, the additional data they want, and how to effectively organize all this data to yield the desired business outcomes” [5]. So data strategy should include business objectives together with the vision of the organization and its technical and organizational capabilities to reach the objectives [18]. In developing any strategy the key element is to define what the goals are that need to be accomplished. For this, companies should start with data strategy statement: objective, scope, and advantage [15]. Strategic objectives explain what this strategy supposed to achieve, scope draws the limitations of data operations for the organization and the advantages are some unique set of values or capabilities that the organization has to deliver a customer value [15]. So, “the strategic statement is a clear statement outlining the organization’s objective, where it will operate, what value it creates for the customer, and what set of activities it will perform to achieve differentiation” [15]. In the case of data strategy, that objective needs to be formulated around data and its value. This strategic statement has to align with the business strategy of the company, has to comply with the mission, vision and the core values, which also creates a standardized pattern for the decision-making across the whole organization.

However, “a recent survey suggests that only 13% of organizations have a formal data strategy that covers managing data and unstructured content” [19]. These numbers are very low keeping in mind that many companies generate a vast amount of data, especially companies that are dealing with online sales and online service. It could take a lot of time to look at all this data and manage it. That is why a data strategy needs to be created, that would prioritize some data assets over the other. Also, there are many tools and technologies on how to manage data and if a company has invested in the wrong technologies that could cost a lot of money and time. To prevent that, data purpose and data objectives need to be addressed in the strategy.

1.2.4 Data Challenges in the organization

Most of the issues regarding data in the organization come from within the company and especially the issue starts from the management “by refusing to create a culture around data

and not prioritizing the proper funding and staffing for data management" [20]. According to the several pieces of research conducted the biggest data challenges for organizations are:

Lack of top management support and engagement

“Executive support is critical for bringing clarity of vision to the overall data strategy and for bringing stakeholders in multiple disciplines and departments together to maximize value creation” [21]. It is much more likely that the project, program or a strategy will be successful if the management is actively engaged in the overall process. Without top management engagement, any initiative could easily end or even would not be executed at all. For the data initiatives, it is even more challenging as data projects are mostly internal and create an overhead, meaning that there not be a tangible value from executing the project.

Poor data quality.

Data quality is one of the top challenges for the organization regarding data and “it is a major obstacle to delivering actionable insight with 50% of companies citing this as a major challenge” [21] from the research conducted by EY. Another research, conducted by Experian, claims that close to 91 percent of the companies are suffering from poor data quality that resides in data errors [22]. The general data errors are related to incomplete or inaccurate data, outdated information, duplicates, inconsistent data, typos, and spelling mistakes and even data that has been entered in the incorrect field in the system [20] [22]. Poor quality has some consequences in an organization as false information (internal and external), inadequate analysis of the customers or other relative business elements, and also a loss of revenue from lacking a clear overview of the business and its processes in the organization [22].

Data integration complexity.

Companies find data integration as a big challenge. The issue is with a complexity that comes from “facilitating the access, transmission, and delivery of data from the numerous sources” [23] and loading that datasets into other different systems. It can become a pain for the company and it costs additional resources and time for the complex data integration and many faults and bugs for the data and system accuracy.

Lack of data specialists

Data specialists are needed to execute the data strategy and create value for the company out of data. However, in the research by EY, 38% of companies are searching the data specialists and 21% of the companies are running internal programs for employee retraining to data savvy specialists [21]. Data experts are needed in all the parts of the business and organizational processes "from building an effective business case and modeling ROI, to

selecting and implementing a successful technology strategy and appropriate, cost-effective architectures" [21]. This is a challenge because data and data related technologies are growing and changing so fast that the gap with the emerge of talents is growing as well.

Privacy and Security

Another big challenge for the companies is data privacy and security. The last couple of years because of the new upcoming European General Data Privacy Regulation, data became a painful topic for all companies, especially for those that are dealing with customer data. It is changing the way that companies were managing their data before, especially in the technical sense as the big part of GDPR is the customer right to ask "to be forgotten". This means to the company that at any given time when the customer request to delete all information about him/her, the company has to delete all the records from all the systems. To make this possible the company needs to know where the data is stored and processed at all times. With the vast amount of data and many different systems for processing and storing it becomes a big technological challenge to execute this kind of request which also requires a lot of financial resources.

1.2.5 Fast-Growing Companies

This subchapter discussed the profile of the startup companies that have a significant growth and what kind of organizational challenges they are facing.

Startups have begun with ideas, grew to be small businesses and if successful have reached the level of a mature enterprise with sustainable growth. According to some articles [24] [25], there are 5 stages of the business lifecycle: 1) Development 2) Startup 3) Growth And Establishment 4) Expansion and 5) Maturity. Fast-growing companies are somewhere between the establishment and expansion stage. They are already out of the struggle to find the first customers as the startups but not at the maturity stage yet, the revenue keeps growing exponentially.

In every stage company has different challenges to deal with, therefore the company has to be flexible and be able to adapt its strategy and objectives with the new stage of the company's lifecycle [25]. With the growth of the company the business goals and priorities changes, as well as the overall strategy of the organization.

The growth of the company is very important, "it remains one of the most important factors in valuation, access to capital, and shareholder return, as well as the ability to attract top talent"

[26]. That is why the majority of companies in software development, that have revenue greater than $100 million per year, are super growers (>50% CARG) or growers (10-50% CARG) [26].
Fast growing and not yet matured companies usually have the criteria of the “Supergrowers at any cost” [26]. Despite the low profits, the companies are investing into the innovative product and try to capture as much market share as possible at this stage. “Many SaaS businesses have pursued this route during their early days and growth stages, using venture financing to fuel negative margins and high burn rates.” [26] The rapid growth is so important to that kind of companies, that they are putting all of the focus only on the direct revenue and innovation.

Unity Technologies is also considered as the fast-growing company with the several million dollar investments in the past few years [27].

1.3 Problem formulation

How to build a comprehensive data strategy in the organization? This problem formulation is the main research question in this Master Thesis. The following are the sub-questions that would help to answer the main research question.

- What components should data strategy include?
- What are the main data challenges in the company?
- How to align data strategy with business strategy and other strategies in the company?
- What are the existing frameworks of data strategy?

1.4 Scope and Limitations

The scope of this research is the creation of the data strategy for the single case study. Although there are some data strategy frameworks already developed, it was chosen to construct a new, more applicable and adaptive framework for the case study. In regards to mass customization, the existing frameworks might not be sufficient enough for the specific company and need to be adapted accordingly. "Mass Customization relates to the ability to provide individually-designed products and services to every customer through high process flexibility and integration” [28]. Even though mass customization describes widely produced products or services, it is no different when talking about the applicability of the standardized data strategy framework for the company. With the concept of customization, data strategy is constructed and tailored for the specific needs of the organization, which would not possible

with a standardized framework. Furthermore, the conceptual framework is used in this Master Thesis to tailor the data strategy towards the case study. Existing data strategy frameworks are too general for the needs of the company and strategy theories and models do not have the sufficient data factor in them. That is why the conceptual framework has been chosen for this research.

The main objectives of this Master Thesis are to research existing data strategy frameworks and strategy theories and to investigate the case study company, have interviews with department senior managers. Existing frameworks and theories will give the base for the conceptual framework and the interviews will formulate the big picture about the current data status and challenges in the company, as well as the perspective about data strategies and possible future plans towards it. This research focuses on the strategies and their elements alignment across the company, data strategy domains and their applicability for a case study, but not the actual implementation of the framework.

There are also some limitations to this Master Thesis, in order to set the boundaries of the research:

- This research does not investigate the financial part of the data strategy, even though, the financial benefits of having a data strategy are briefly mentioned.
- The technical part of the data strategy is discussed only on the conceptual side due to the main focus on the strategy alignment part of the research.
- The analysis of this research is based on the creation of the data strategy framework but not its implementation.

2. Methodology

This chapter focuses on the research methodology for the Master Thesis, describing how the research has been conducted to answer the main problem formulation. Throughout the research, a literature review of strategy theories and frameworks has been used as well as state of the art data strategy frameworks has been reviewed and investigated. Exploratory and qualitative research methods have been used to collect the data.
The Master Thesis research is also based on the single case study that the problem formulation question is pointed at and will help to answer it. In this section, the methodology framework will be presented to show the research flow to answer the problem formulation.

2.1 Research design

This section explains what research designs are used in the Master Thesis and the methodologies to fulfill those designs.

2.1.1 Exploratory Research

The thesis is based on the exploratory research, which is designed to conduct the research that has not been clearly investigated before and does not have a defined outcome [29]. The value of exploratory research is to assess the phenomena from a new angle by asking questions and discover new insights [30]. The exploratory design is used when the problem is not certain or if it was not focused on before. There are several ways of conducting this type of research [30]:

- Literature review
- Interviewing subject matter experts in the field
- Focus groups interviews

The advantages of exploratory research are that it is very flexible and adaptable to the changes, the research focus might change from being too extensive in the beginning to a narrowed down [30].

The exploratory research design is applicable to this Master Thesis due to the topic uncertainty and ignorance by the companies, as only 9 percentage [3] of companies are able to use their

data in the very effective way that would bring them value and benefits, also data strategy is a relatively new subject for organizations.
Exploratory research has been used to collect data by the interviews within the company, selective research literature review, and theories.

2.1.2 Descriptive Research

Master Thesis also uses descriptive research design. By the suggestion of the name, the descriptive design is focusing on describing a clear picture of phenomena. It is used sometimes as the extension of the exploratory research [30].
In this Master Thesis, the descriptive design is used to help to understand the ground of the strategy, theories for strategic alignment and data strategy frameworks.

2.2 Research strategy: Case study

A case study is a research strategy that includes empirical data collection of the specific case with the real-life context [30]. A case study gives the opportunity to observe, explore and investigate the phenomena in its natural occurrence and construct a clear understanding of the problem [30]. A case study is commonly used when the research subject is not well examined, not well understood or is particularly new [31]. The study case uses interviews, observations, documentary analysis and questionnaires to make research. Study case is also a part of the exploratory research that is used in this Master Thesis.
A single case study of Unity Technologies is used in this research. Single case study represents a specific case or a group of specific cases that could be a company, organization or any specific single group [30]. The advantages of using this type of study are that a single case study gives an opportunity to have a systematic approach of collecting data, analyzing data and observing the environment over a long period of time, which gives in-depth insights to the case study [31]. It also gives an opportunity for an easier internal validation or testing of the research applicability to the case. A single case study could stand for a representative or a typical case as well [31], which in this Master Thesis it stands for fast-growing companies like Unity Technologies that are generating and using huge amounts of data.

2.2.1 Theories and Literature Review

Theories for strategic alignment has been researched and investigated in this Master Thesis to have a theoretical background that could support the data strategy framework.

In the literature review part, different data strategy frameworks, data strategy domains, data maturity assessment tools and the concept of data as an asset have been discussed to build a strong overview of the existing frameworks and methods in the problem area.

2.2.2 Interviews

Interviews have been used due to the exploratory nature of the research and the research strategy of the case study. It is also recommended in the exploratory research design, that the subject matter expertise should be interviewed. The respondents are the experts in data and data processing, senior managers that are experts in their own department strategy and also have a big picture of an overall organizational strategy.

The first two interviews were taken with the general manager of the data warehouse team. The interviews were not structured and informal conversation type. In the first interview, there were no predetermined questions, it was more like a general discussion on the subject matter to get the perspective and the ideas on the topic which helped to form the way of this Master Thesis. Before the second interview, questions were sent to the DWH manager on the email. The questions have been formed according to the first informal interview where the topics have been touched about the data related challenges in organization, data strategy, data architecture, the scope of useful data, data processing tools, data quality, data monetization and the future strategies for Unity Technologies. In the interviews, DWH manager has addressed the challenges that the organization had regarding data and the position of the data strategy and strategy alignment in the company.

After the better view of the issues regarding data and strategies in the organization, there were several semi-structured interviews with senior managers at the company. The questions were prepared in advance and asked during the interview, however, depending on where the interview was happening there were some additional questions asked or some questions skipped from the questionnaire list. The selected interviewees were the general managers of their departments so that they would be able to answer the questions about the organizational strategy and strategy in their own departments.
The main objectives of these interviews were to investigate what are the current issues with data, whether there is an alignment between business, IT and data science departments and to find the department manager's perspective on the necessity of the strategy alignment. The questionnaire (Appendix A) has been designed according to the literature review on strategic alignment and data strategy frameworks, as well as the similar research [32].

Six interviews were conducted in total with the departments of DWH, Business Analytics, IT and integration systems, Data Science. All of those interviews gave a perspective on the study case and shaped the data strategy framework and guidelines for Unity Technologies. The main findings from the interviews will be presented in the empirical data chapter.

2.2.3 Validity Workshop

As one objective part of this Master thesis is to develop a data strategy framework for the study case, it needs to be evaluated by the company that the framework has been built for. The validation of the framework will be taken in the form of the workshop with the DWH manager, who is the initiator of the idea to create coherent a data strategy across the company.

Although some researchers have argued that there is no need for the validation of qualitative research, they have later concluded that there is actually a need for some kind of validation check, quality check or measure of their research [33].

The validity of the case study is important because it determines if the findings truly measure what the researcher intended to measure and how applicable or truthful the results are for the case study [33]. In the qualitative research, validation of the findings checks if those findings are true and certain, that means if the outcomes of the research are as accurate and reflect the case and if they are backed up by the credible sources [34]. Validation of the qualitative research also creates the confidence and reliability of the findings [33].

2.2.4 Triangulation

The research is using data triangulation for data collection. Data triangulation is the method that use multiple data sources to make research more valid and answer the research question from multiple perspectives [34].

Triangulation is a common method for case studies, it refers to using different data gathering techniques and different sources of data that are used to support the research findings in the study case [30]. Triangulation in this research is reflected by using interviews and literature review for the data strategy frameworks and theories to support the case study.

The main advantages of the triangulation method include "increasing confidence in research data, creating innovative ways of understanding a phenomenon, revealing unique findings, challenging or integrating theories, and providing a clearer understanding of the problem" [35]. Thus, using literature review, interviews and validation of the analysis adds the depth to the findings of the research that would not be possible only with one method [34].

The downside of the triangulation is that requires more time and resources to collect data in different methods, as well as the possibility of the collected data being conflicting and bias towards research and interviews respondents [35].

2.3 Theoretical framework

A theoretical framework is used to help to answer a problem formulation that has been applied to the study case. After the review of the strategy theories, there was a need for the comprehensive theoretical framework that could fit the study case and includes different parts from different theories. One of the ways to make it happen is to use the conceptual framework. The whole conceptual framework has been developed and applied to the case study. It includes the mix of strategy alignment theories and the most suitable parts from data strategy. The whole conceptual framework will be discussed in the upcoming chapter "Conceptual Framework Analysis".

2.4 Methodology framework

This is a graphical view of the methodology of how this research has been conducted and the main question of problem formulation has been answered.

The initial project idea for the Master Thesis has been chosen after reviewing literature and trends in the ICT industry. As the topic of data and its management is pretty broad, after the primary literature review and informal interview with the subject matter expert, the initial idea has been broken down to a more specific topic of data strategy and the first version of problem formulation has been established. To answer the problem formulation, a research method of a single case study has been selected. The problem formulation has been reformulated several times after the data collection and better understanding of the problem area.
Data collection for the research includes:

- Theories and literature review: strategy alignment theories, literature review of data strategy frameworks and domains, data assessment tools and data as an asset concept;
- Empirical data: interviews with the case study company and the validation workshop of the conceptual framework;

Data collection helped to create a more clear view of the existing solutions for the problem formulation and gain the insights into the case study (company profile, the current status of the company, challenges with data, view on data strategy).

Also, after the review and validation of the conceptual data strategy framework, the framework has been amended to fit the case study.

The researcher continues with developing a conceptual framework for data strategy combining theories and literature review with empirical data. The conceptual framework is then analysed in details for the case study of Unity Technologies.
Finally, the conclusion is made that helps to summarize the research and pinpoint the answers of problem formulation.

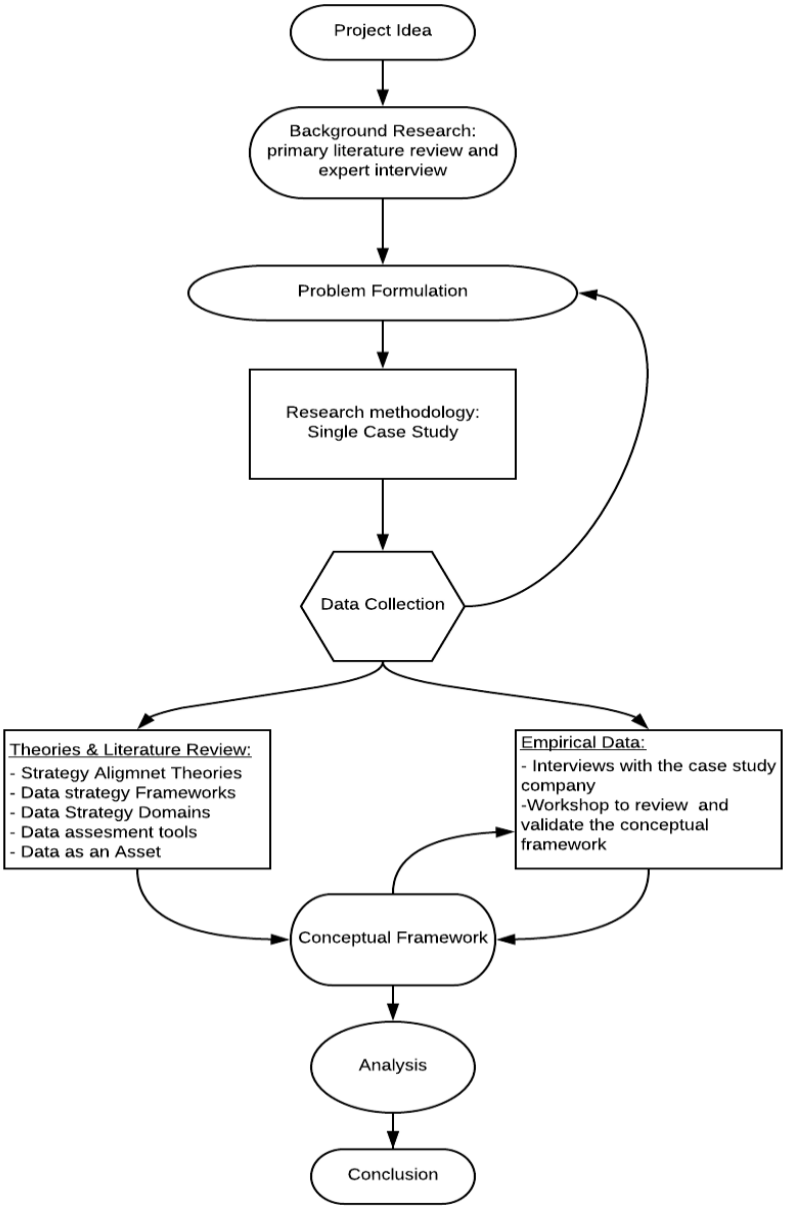

Figure 2. Methodology Framework

2.5 Summary

Methodology chapter explains the whole workflow of the research, how the problem formulation was formed and the process of how it was answered. The chapter also explains what methods have been used and why.

3. Theories

For the strategy to be successful it needs to be aligned with other strategies in the organization. “Well-crafted strategies are roadmaps to places that yield a competitive advantage and generate value for the firm” [36]. To craft a good strategy in business it requires to have a theory behind it, as in any scientific area [36]. Also, there is a lot of division and miscommunication between the IT strategy and the business strategy in the organization which requires to have a good framework to solve it [37].

This chapter discusses theories, models and a conceptual framework that would contribute to the theoretical background of the study. “A theory is a set of interrelated concepts that provides a systematic view of a phenomenon for the purpose of explaining or predicting” [38]. Meanwhile, a model is a simplified version of a theory usually expressed in a graphical way [39]. Scientists use models to represent aspects of the world for specific purposes [40]. A conceptual framework is a synthesis that comes from a variety of "conceptual or theoretical perspectives" [41]. The purpose of this chapter is to investigate relevant theories and models for later application to the conceptual framework.

3.1 The Three Levels of Strategy

A strategic approach could be used in all parts of the organizational life cycle, from the most generic overall strategy to the specific, short-term goal-oriented strategy [42]. The strategy is broken down into three major levels: corporate strategy, business strategy and functional strategy [43]. Those three levels of the strategy interrelate with each other according to the organizational structure from the CEO of the company to the managers of each department as shown in Figure 3.

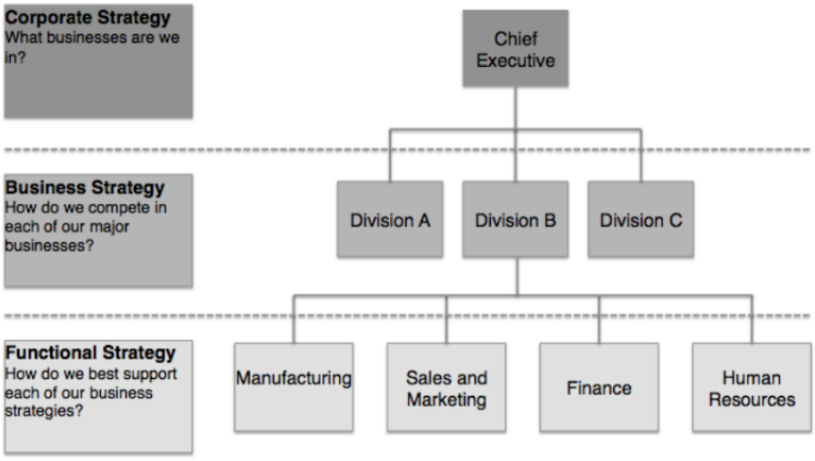

Figure 3. Three Levels of Strategy [43]

The corporate level strategy is the core and the foundation of the company, that directs the way the company is going and what it wants to achieve. Then the business strategy follows which indicates how the goals that were created as the corporate strategy could be achieved and how the company is going to play in the market competition. And then finally, there is a functional strategy that supports each business objective defined earlier, whether it might be IT strategy, Sales Strategy or any other strategy needed to support business objectives [43]. The ones responsible and driving the strategies are the managers of the departments or initiatives, they are making all the decisions behind the functional strategies but also they need to make sure that it aligns with the business and overall corporate strategy. The functional strategy level is crucial for having an overall successful strategy, that also involves the alignment and interconnection between all three levels [43].

3.2 SAM Model

Strategic alignment theory has been researched by several people at different years. It was relegated from the 90s to this day. The below table shows some of the definitions of strategic alignment in the organization:

Definitions of Strategic Alignment	
Definition	**Source**
Strategic Alignment Model describes "strategic fit (the interrelationships between external and Internal components) and Functional Integration (integration between business and functional domains)".	“Strategic alignment: Leveraging information technology for transforming organizations”, Henderson and Venkatraman, 1993,
“…the degree to which the information technology mission, objectives, and plans support and are supported by the business mission, objectives, and plans.”	“Factors That Influence the Social Dimension of Alignment between Business and Information Technology Objectives”, Reich and Benbasat, 2000, p. 82
“Strategic alignment is a mechanism by which an organization can visualize the relationship between its business processes and strategies”	“Strategic Alignment of Business Processes”, Evan D. Morrison, Aditya K. Ghose, 2011
“strategic alignment can be seen as an array of bounded choices made in order to resolve strategic ambiguity”.	"Strategic alignment: a practitioner's perspective", Journal of Enterprise Information Management, Bruce Campbell, Robert Kay, 2005
“Strategy alignment addresses “how IT is in harmony with the business, and how the business should/could be in harmony with IT. Alignment evolves into a relationship where IT and business adapt their strategies together".	“Assessing Business-IT alignment Maturity”, Jerry N. Luftman, 2001

Table 1. Definitions of Strategic Alignment [44]

As the researchers were debating the alignment of strategies in the organization, most of those descriptions have something in common, they are talking about integration between IT, business and organizational strategy. However, Strategic Alignment Model [43] (SAM) developed by Henderson and Venkatraman already in the 90s is still used for today's concept and it is still one of the most used tools in business and IT alignment that most researches and business experts apply for the strategy alignment process [45] [46].

Strategic Alignment Model has been developed in 1993 and it was the first model to define the interconnections between IT and business strategies. However, because of its conceptual status, the model is still applicable to this day.

SAM is a comprehensive approach model that visualize an organization and its IT department with internal and external components. Internal level covers the alignment between IT and business infrastructure and external level includes IT and business strategies alignment [47]. SAM model that is shown in Figure 4., also describes "the cross-domain link which encompasses the relationships between IT strategy and business infrastructure and vice versa" [47]. There are 3 types of integration between IT and business domains in the model [47]:

- Intellectual alignment / strategic integration: this is the alignment between IT and business strategies which indicates the support of both strategies mission, objectives and main milestones [47].
- Operational alignment/integration: this is the alignment "between organizational and IT infrastructures and operations and denotes the integration between organizational policies, procedures, and systems with IT architectures and processes" [47]
- Cross-domain alignment/integration: it describes "the degree of fit and integration among business strategy, IT strategy, business infrastructure, and IT infrastructure" [47]. It also addresses the multiple choice of relationships between all of the components in IT and business scope [47].

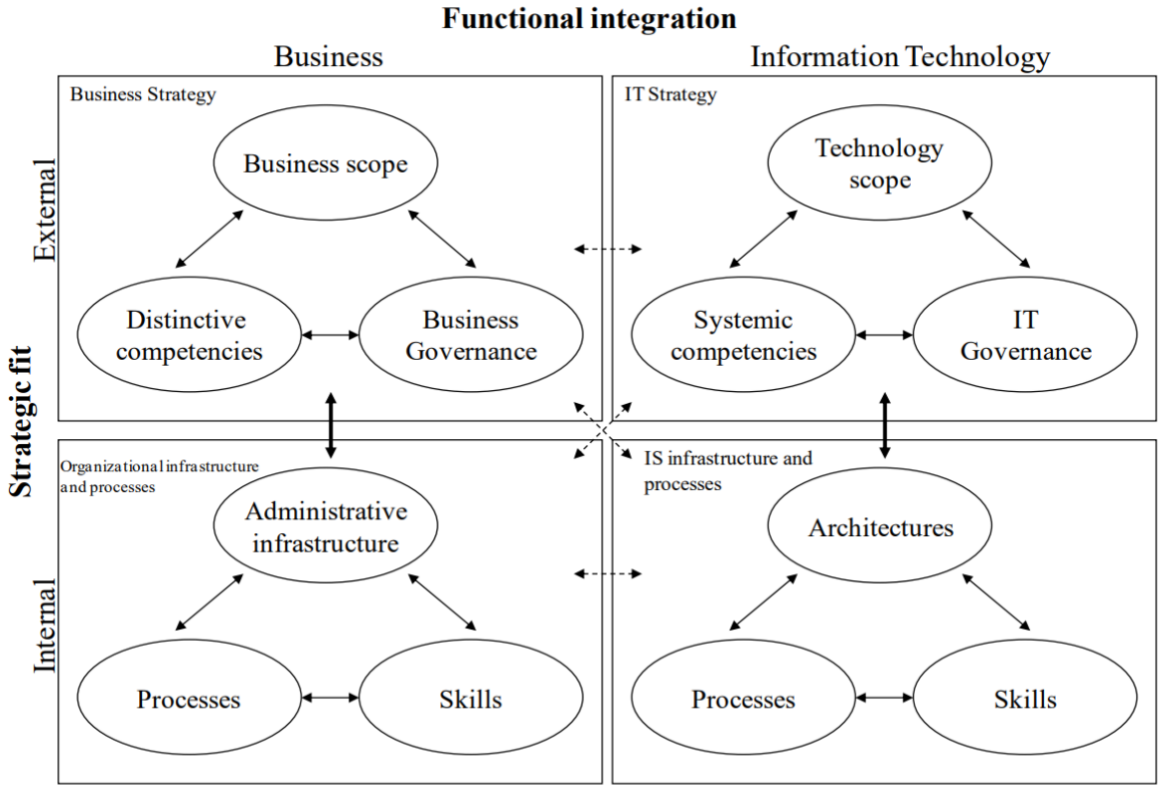

Figure 4. The Strategic Alignment Model [48]

The main purpose of this model is to guide companies to reach the strategic alignment between different domains and create a comprehensive strategy that works as a coherent solution.

3.3 Co-evolutionary IS/Business alignment Model

Co-evolutionary IS alignment Model is based on the Henderson & Venkatraman SAM model and some other strategy alignment theories that followed after [49]. Co-evolutionary IS/Business alignment theory suggests that it can provide critical insights for managing the IS/Business alignment which has a big complexity and many different levels. Co-evolutionary alignment framework addresses the "multilevel aspects can be seen where IS should be, on the one hand, aligned with user's needs (Individual level), between Business and IS departments (Operational level), and between Business and IS strategies (Strategic level), and on the other, adapted to the changing external environment" [50].

The Coevolutionary IS Alignment model (Figure 5.) illustrates the three levels of the framework and the interrelationship between them. There the three levels are: "(1) Strategic level – coevolving IS and Business Strategies; (2) Operational level – coevolving Business and IS

domains; and (3) Individual level – coevolving IS infrastructure with users needs" [50]. Integration between several levels in the organization, between IS and all the organizational objectives, creates the comprehensive understanding that those several levels of alignment increase organizational performance, business outcomes and competitive advantage [51].

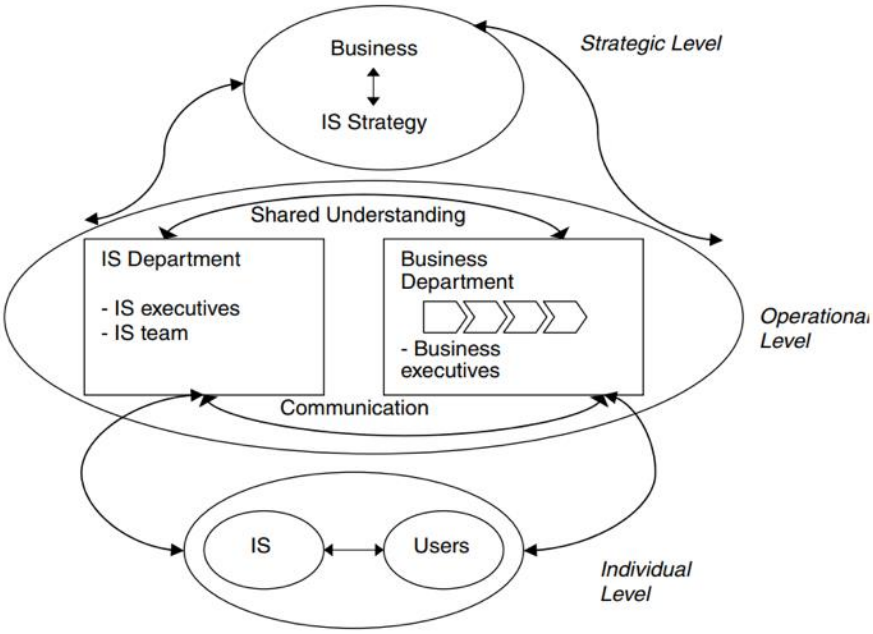

Figure 5. Coevolutionary IS Alignment [50]

3.4 Conceptual Framework

The previous paragraphs discuss three theories that are related to strategies, they have some different perspectives on the subject matter. Those all three theories are valid and gives a value for the research, however, the theories have been developed before the data subject became an important matter for the companies. It would be hard to apply the full theory to analyze a case study due to the different company profiles (the concept "one size fits all" is not applicable when talking about individual case studies) and also due to the fact that the market environment is changing so rapidly that some parts of the theories get obsolete as it lacks a new perspective. Nevertheless, theories give a good ground for the research and the most efficient way to use them would be to combine them all using a conceptual framework.

A conceptual framework is a framework combined with several different theories, models, approaches or concepts and is usually realized in the graphical or a written form [52].
A conceptual framework can also be a theory but a tentative one or incomplete, that could sometimes be compared to a research problem as it is also not defined accurately in the

beginning of the research but rather gets developed in the process when there is more information about it [52]. Therefore, a conceptual framework is something that is constructed on the way with the flow of the research. “It incorporates pieces that are borrowed from elsewhere, but the structure, the overall coherence, is something that you build, not something that exists ready-made” [52]. Every theory used needs to be carefully analyzed and the valid points need to be put together to construct a concept that would be adequate and fit the case study.
When using the conceptual framework, it gives the researcher a possibility to look at several different perspectives “that combines divergent mental models to expand and deepen, rather than simply confirm, one’s understanding” [52].

The conceptual framework will be used and it will be presented in the analysis part of this Master Thesis. It will be used to create a combined data strategy framework for the Unity Technologies case study.

3.5 Summary

The theories chapter discussed strategic alignment theories to have a theoretical support for the data strategy framework. All the theories conclude one common element, the alignment between business strategy and other strategies in the organization. These theories will be combined through the conceptual model and implemented into the data strategy framework for the case study.

4. Literature Review

In the literature review chapter, the state of the art data strategy frameworks will be discussed as well as data strategy domains, data maturity assessment tools, and data monetization approaches.
A state-of-the-art represents the most recent researches, technologies, tools or ideas in an area of concern [53]. It includes the emerging trends and priorities, as well as the literature produced in the past decade [53].

4.1 Data Strategy Frameworks

In this chapter, several of the data strategy frameworks will be discussed. Data strategy frameworks guide companies to the best way of how to initiate data strategy in the company and implement it. However, the frameworks that will be discussed do not include the actual parts of the data strategy, only the methods of how to reach the best results in creating and maintaining one. Below are the frameworks that have been selected from various research papers.

4.1.1 ThotWave's Data Strategy Framework

ThotWave's Data Strategy Framework [15] has been developed for the ThotWave technologies specifically, but could also be applied for the other organizations. This framework consists of six stages written below and mostly talks about the linkage between the data strategy and its execution. This framework also is based on the design thinking model which is a user-centric process [15]. In this case, all the users are the stakeholders of the data strategy so it is being framed around them. ThotWave's Data Strategy Framework also uses the agile method to develop and execute the strategy, and the stages of the process are iterative and flexible to the changes [15].

1. Develop the strategy

In the first step of developing the data strategy most of the work goes to the data or analytics organization in the company, or if there are no such departments then the data committee should be established. The main task for this data organization is to address all current and possible issues that a company faces in regards to data. Design thinking model should be applied here to create all the possible user stories and find out all the possible scenarios for that usage, from which the organization would be able to address the main issues and start to

develop data mission and vision in the company together with the strategic objectives. Also, in this step, the measurements of success needs to be developed that also fall under the main KPIs of the entire company.
After all the above has been done, the data organization is able to create a data strategy statement that lays out data purpose, data scope, and data advantage.

2. Create the Roadmap

In this step, the data organization needs to create a long-term execution plan or a data strategy roadmap. Creating a roadmap includes describing a specific course of action, milestones/deliverables, timelines, resources, and budget. Risks and their mitigation plan should also be considered in this step.

3. Align the Organization

In this step, a new data strategy needs to be aligned with the strategies of the other impacted business units in the company. All the stakeholders that are impacted by the new data strategy need to involved in the creation and evaluation process of the data strategy. That also means, that all the impacted business units in the organization have to align their strategies, i.e. modify their objectives according to the new data strategy. The outcome of this stage could be a stakeholder management plan and a change management plan.

4. Execute the Plan

In this step, data organization has to ensure that this long-term data strategy is linked with everyday operational work. This means that the data strategy needs to be implemented in the everyday workflow of all the business units that are impacted by data. A data governance plan needs to be developed to ensure the successful implementation of the new strategy.

5. Monitor and Learn

In this step, the measurements of success are applied for monitoring the execution of the data strategy in all of the business units. For an effective monitoring, a data management meeting is recommended where the results are being reviewed and evaluated.

6. Test and Adapt

This stage is very closely related to the previous stage "Monitor and Learn". The insights from the stage 5 are analyzed and some new changes could be applied to the data strategy, then tested on the internal data and if results are successfully adapted to the whole scope of data strategy.

4.1.2 Three Tiers Data Strategy

The students from Samford University and the University of Connecticut have developed Three Tiers Data Strategy. This data strategy framework [54] has been developed under the assumption that big data reaches the limit of the organization's capabilities to manage it [54]. Framework treats the data as it is a success factor for the organization to strive upon. Three-tier data strategy framework includes usability of data, technical tools, infrastructure and digital transformation of business.

Table 2. illustrates each tier and its adoption in the company:

Table 1. Three tiers of value creation due to the big data phenomenon

Data as a tool	Managers are able to solve traditional value chain problems more efficiently and effectively; existing capabilities are improved through real-time, customized decision making for individual consumers
Data as an industry	Spin-offs and new ventures are created to specialize in acquisition, storage, and analysis of data, construction of infrastructure, and development of software devoted to handling big data
Data as a strategy	Visionary leaders develop companies dedicated to building data resources to allow them to develop radically innovative business models that wed traditional and modern strategic thought

Table 2. Three tiers of value creation due to to the big data [54]

Tier one: Data As A Tool

This is a basic tier for most of the traditional companies. Data is used to solve everyday business issues and make business operations more efficient and effective. By extracting data and making it meaningful with analytical tools data manager take data-based decisions that improve business operations like "product development, marketing, sales, distribution, customer service, and other traditional value chain activities" [54]. These data-based decisions create value mostly for the individual customer by making his or her product more customized, by creating individual based advertising and suggestions for incremental products.

Tier Two: Data As An Industry

These tier companies are focusing on internal data infrastructure and all the resources that are related to it. The newly established companies are developing the infrastructure and technical capabilities on their own and the more mature companies are hiring 3rd party providers to create the infrastructure for them. It is applicable for the companies that the main business model comes from data or some of the data is monetized.

Tier Three: Data As A Strategy

This tier talks about changing company strategy to the innovative data-oriented business strategy. Data should be focused as a central core in the organization, moreover, the focus should be on data flows and integrations [55]. The whole ecosystem of the company should be based on the data from their customers, systems, and products. This approach is a combination of tier one and tier two and it is usually adoptable by the companies that are able to invest in the data strategy development without getting an instant revenue from it, just creating a value in the data itself for the long-term goal.

4.1.3 Harvard Data Strategy Framework

The authors of this framework claim that it could be applied across many different industries and to the companies with different profiles and stages of data maturity in the organization. The framework is based mostly on the experience with the big global insurer corporation AIG and other large enterprises where some elements of this frameworks have been implemented. This framework is distinguishing two key areas of data strategy: Defense and Offense. They are described by different business objectives and actions to achieve them.

- Data defense's primary purpose is to control and govern the data and data flows. Activities of data defense "include ensuring compliance with regulations (such as rules governing data privacy and the integrity of financial reports), using analytics to detect and limit fraud, and building systems to prevent theft" [4]. Also, defense activities help to keep or create data integrity "by identifying, standardizing, and governing authoritative data sources, such as fundamental customer and supplier information or sales data, in a "single source of truth" [4].
- Data offense, on the other hand, is focusing on the business objectives that are related to increasing revenue and profit, creating the value for the internal organization and for the customers. Data offense mostly includes the activities that are related to customer data analytics, modeling their client behavior and customizing products, also includes data about the market which helps to make data-oriented managerial decisions.

Every organization needs both offensive and defensive data strategy mechanisms to balance out the protection and control and creating a value out of data. The challenge here is to create the right balance between defense and offense in the company, between standardizing the data and making it flexible, which usually depends on the industry the company is in. If the industry is highly regulated then the defense side would take over, but if the industry is more competitive, open and customer-oriented then the offensive strategy is more appropriate.

The defense and offense conception is also illustrated by the approaches of "single source of truth" (SSOT) or "multiple versions of the truth" (MVOT).

- SSOT is a primary source of data that is unchangeable and is vital to deliver and run business operations. "Customer details, supplier details, and product information should come from an SSOT, which provides an excellent level of control, perfect for those who are concerned with regulation and compliance" [56]. SSOT data makes organizations structure less chaotic and makes sure that there is a standard of data sources in places that are needed. For example, SSOT means that in all the company's systems a specific customer "ABC" has to be named the same and not have different listings as ABC Corp or ABC Inc. This single source prevents the company and different departments from confusion, misunderstandings, and mistakes.
- The MVOT concept talks about transforming the data from SSOT into multiple versions of information, making this standardized raw data into the meaningful, flexible and customized piece of information. The idea is that each department in the organization adopts SSOT data by their needs, relevance, and requirements.

The main idea of this framework is that “It requires flexible data and information architectures that permit both single and multiple versions of the truth to support a defensive-offensive approach to data strategy” [4]. This approach is illustrated in Table 3.

	DEFENSE	OFFENSE
KEY OBJECTIVES	Ensure data security, privacy, integrity, quality, regulatory compliance, and governance	Improve competitive position and profitability
CORE ACTIVITIES	Optimize data extraction, standardization, storage, and access	Optimize data analytics, modeling, visualization, transformation, and enrichment
DATA-MANAGEMENT ORIENTATION	Control	Flexibility
ENABLING ARCHITECTURE	SSOT (Single source of truth)	MVOTs (Multiple versions of the truth)
FROM “WHAT’S YOUR DATA STRATEGY?” BY LEANDRO DALLEMULE AND THOMAS H. DAVENPORT, MAY–JUNE 2017 © HBR.ORG		

Table 3. Elements of Data Strategy [4]

4.1.4 Eckerson Ten-Step Process

This data strategy framework is created by Eckerson Group [57] which is the leading in the research and consulting business of data science with more than 20 years of experience in the field. Eckerson consultants developed this framework [10] for companies to evaluate and

improve their existing data strategy or to assist in creating a new one. The framework is simple enough, easily understandable and as the title suggests consists of 10 steps [58]:

1. Build awareness:

   The initial step of the framework is about drawing the top management's attention and getting the support or even the engagement from them. The step also includes establishing a data organization or a data program in the company as a starting point.

2. Assemble a team:

   When the initiative of data program has been established, it needs to have a team of all the different specialists and people of interest across the company. The best way is to create “a cross-functional team of senior-level managers who have a vested interest in improving the quality of data and analytics at the organization” [58]. A cross-functional team is needed so that the strategy would have inputs from all the different sides in the company and later it would be easier to implement in the strategy into different departments as well.

3. Educate the team:

   As usual, when introducing something new, employees need to be educated. For the data strategy initiative, they need to be educated about data, data strategy and its purpose [58]. Furthermore, the newly created team needs to be trained with the knowledge of data expertise from the inside or outside consultants which would know the latest trends in the industry.

4. Assess the current state:

   This step is for evaluating the current situation of the data status in the company. What data is being collected, how it is being collected, which data is used and etc. The assessment could be done internally by using some of the tools like SWOT analysis or the external consultant could be hired to conduct an official and industry approved data assessment [58]. Both ways, the assessment of the current state of the data is crucial, without knowing where does the company stand it would be really hard to build the way forward [58].

5. Develop the vision:

   The assembled cross-functional team should come up with the future vision, the future state of the data strategy in the company. It also needs to be aligned with the business vision and overall strategy of the organization [58]. The vision should not be vague, it should fulfill the operational needs and improvements and should create a data statement which then should be used as a cornerstone for the whole data strategy [58].

6. Develop recommendations:

   This step covers the creation of data objectives and the guidelines of how those objectives should be attained. The objectives are the high-level goals of data strategy

that support the data strategy vision. The objectives also should be constructed as a clear action list for achieving the data strategy vision, which should also be coherent with the objectives of the business strategy [58].

7. Develop a roadmap:

   This is a high-level plan for the team to develop. Usually, the roadmap is set up for a few years with the most of the milestones in the first and second year. The dependencies, potential resources, and the timelines are also indicated in the roadmap. This document is similar to the project plan just in the much higher level.

8. Develop a business case:

   This step is for developing a proposal or a project plan that would be submitted to the board of executives for the approval. The proposal should detail the actions and "resources required to implement the strategy, including capital investments, new processes, new hires, and new organizational structures" [59]. It is very important to engage few people from top management in the development of the data strategy so that this program would have a support in front of the board of executives as the proposed change would cost a large amount of budget and would impact all the business areas in the company.

9. Prepare for the change:

   As this is a big change for the organization, it could be really hard to convince the top management and get the approval, however, it could be even harder to change the current state of the organization's culture and processes towards data. Some of the new roles could be required, the business processes could be changed, everyday tasks for some employees could be changed as well. To manage this level of change, it is recommended to establish a change management team, that would create the communication, education and marketing plan for the new data strategy [58].

10. Execute the strategy:

    This is the final step towards successful data strategy implementation. Many organizations that take care of their data also hire a chief data officer (CDO) to oversee all the data areas "including data warehousing, business intelligence, analytics, data infrastructure, and data governance" [59]. Gartner research states that the ideal situation for CDO in the organization chart is to report directly to the CEO and not CIO as it is usually perceived [60]. By reporting to CEO data strategy would cover not only the IT department but the whole business in the company.

4.1.5 SAVV-BIGD Framework

As big data has become a vital part of every organization, data strategy had to be put in place for the alignment. Mike Lakoju and Alan Serrano have developed a big data framework based on the Coevolutionary IS/Business alignment Model to include data strategy into the organization's overview and to show the value of big data [61]. The approach of the framework "suggests that a proper alignment between the Business Strategy of an Organisation and the Big Data strategy is required to visualize the value of Big Data" [61].

The authors of the framework have analyzed two different big data implementation methodologies and combined them with the strategy alignment theories to develop SAVV-BIGB Framework (Figure 6). The framework fills in the gaps between business strategy and big data implementation creating a data strategy framework for the organization [61]. It addresses the alignment between organizational/business objectives and big data objectives in the company.

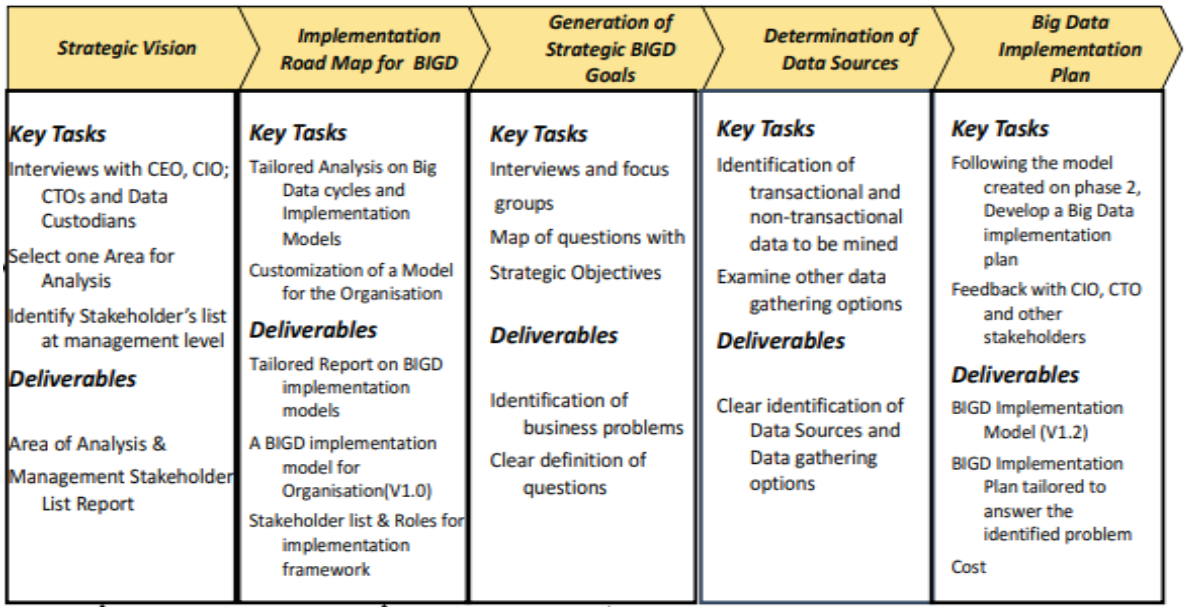

Figure 6. SAVV-BIGD Framework [61]

Strategic vision

This step addresses the alignment between IS strategy, business strategy, and the data strategy objectives. There are many interviews held in this stage to find out the objectives of the strategies and identify the key stakeholders.

Implementation Road Map for BIGD phase

This step involves aligning IS domains with business domains and based on the identified outcome, Big Data implementation model is customized for the organization according to its needs [61]. The specific stakeholders' list is selected for the actual implementation. It is a very important step in the process as it involves analysis of the data within the company.

Generation of Strategic Big BIGD Goals Phase

In this step, there is a continuation of aligning IS structure with business structure and its objectives by conducting interviews, workshops and focus groups. This is also a step where a clear business problem and big data goals are formulated together with the questions related to it. “This entire phase is in line with the operational dimension of the coevolutionary IS alignment model” [61].

Determination of Data Sources Phase

This step is describing the IS infrastructure and the user needs, the alignment between identified business problem and all the data resources, trying to examine all possible data types, sources and gathering techniques[61]. The outcome of this step is a list of approved data sources and data gathering instruments that will be used in the organization.

Big Data Implementation Plan Phase

In this step, the big data implementation methodology is finalized according to the organizational needs identified business problem [61]. A detailed implementation plan is created and validated by the key stakeholders of the implementation. Big data implementation plan has to reflect the identified business problem and fit into the cost that is also set in this phase of the framework.

4.2 Data Strategy Domains

In this subchapter the different parts of data strategy will be discussed, that later on would lie down as the as the data strategy roadmap or action plan for data initiatives. It is important to discuss what should be included in the data strategy. Mike Fleckenstein and Lorraine Fellows in Modern Data Strategy [5] suggest that process, technology, and people are the most important elements in the data strategy. To detail those elements, data strategy domains should be discussed. According to the combined literature review from the academic papers and consultancy companies [5] [15], the main domains of data strategy will be discussed in this subchapter.

4.2.1 Scope

In this part, it is presented what kind of data should be included in the data strategy and what kind of data company can get rid of. First of all, this distinction is needed for the company to have a clear strategy on which data should be focused on and put all the resources to it, so to have a clear data scope and data objective of each data category [15]. Then the company also needs to figure out what data is not valuable and take it out of scope or reduce its storage as it requires a lot of resources and money, and resources usually are limited.
It is important to define the scope of the data that will be included into the data strategy as this amount of data needs to be managed, governed and monitored, meaning that it will require technical, human and financial resources [15]. There are two types of data according to Thotwave [15], the data that organization needs in order to run the business and the data that could improve the business or help to reach certain strategic goals.
The scope of data should be build together with the purpose or objective of that data category, meaning that data should be categorized into groups according to its purpose, the value that it is bringing and objectives that it wants to achieve.
For some categories of data, the value could be also realized into the financial aspect, meaning the purpose of that data is to make a profit from it. By making the categories and aligning with the objectives it helps the organization to spot which data is worth having and which data can be placed in out of scope category and not be collected, managed and stored [15].
An organization needs to create a scope because "successfully managing data consumes resources and there is a limited supply of resources for any organization" [15].

4.2.2 Data architecture and integration

Creating a data architecture means basically to project all the data flow, storage, and usage through the entire lifecycle of the company [5]. Developing data architecture organization should rely on the past data results, storage, data flows, to design a working data architecture and not create it from scratch [15]. Usually, data architecture is developed already for the companies that have been in the business for quite some time, it is hard to find an enterprise that has a finely defined data architecture due to it is resource and time consuming and requires a whole picture that just established companies might not have. So the data architecture in the new companies is being built gradually and not always compliant with all the parts within the organization. That is why at some point when the company is mature enough and ready to take care of their own data, a new data architecture needs to be built.

What are the essential parts of data architecture? Mike Fleckenstein and Lorraine Fellows in their publication of “Modern Data Strategy” [5] provide the structure of data architecture that consists of 5 parts: business glossary, data asset inventory, data standards, data models, data lifecycle diagrams. It is one of the ways how to structure the data architecture and was chosen to investigate as it has both the business management aspect of the architecture and the technical as well.

4.2.2.1 Data Lakes

A traditional data architecture is not able to meet all the requirements and needs of the data, so more and more companies are turning to the new, highly scalable technologies as the data lakes [62].

Data lake is the core part of the data architecture that leads to the data architecture-related decisions and shapes the design of an enterprise data architecture [63]. Data lake is a massive, centralized and highly scalable data repository for raw data (unstructured and structured) of any formats and a data processing engine as well [64]. “The data lake consolidates an organization’s data into a governed and well-managed environment that supports both analytics development and production workloads” [65]. It grasps the functionality of multiple data platforms as databases and data clusters, data warehouses and analytical tools.

Data Lake should have integrations with the rests of the enterprise applications to move further in the data processing and analyzing as shown in Figure 7.

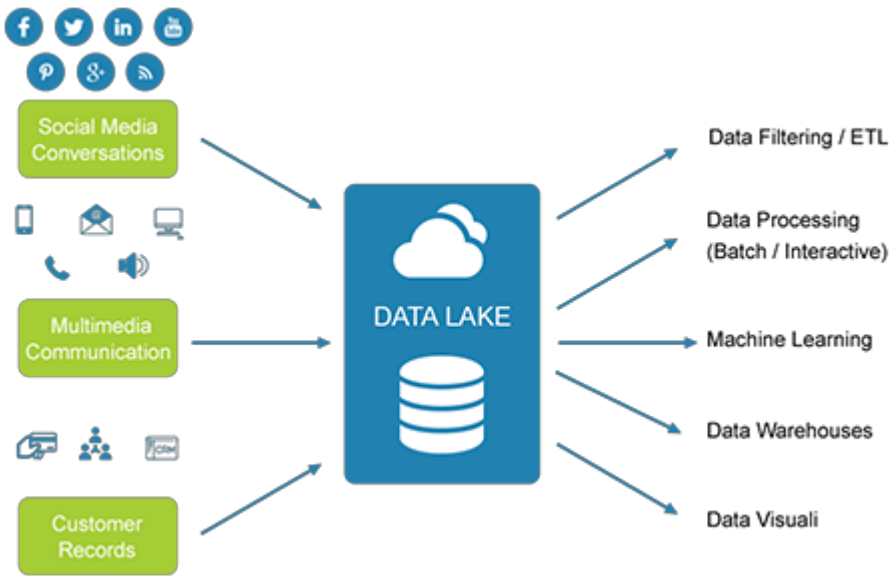

Figure 7. Data Lake Diagram [66]

Capabilities of the data lake [64]:

- High scalability and availability that grows with data
- Data transformation tracking - logging of the data transformation processes, so it is feasible to trace back to the data sources.
- Data classification by data type, data usage and data security [63]

Data Lake is a great technology that can improve the data architecture in the company and all the other data initiatives. However, only if closely governed data lake can succeed, otherwise it could turn to a data "swamp", a chaotic and hardly usable data repository [62].

4.2.2.2 Business Glossary

As data architecture covers all data, data storage and data flows across the whole organization, it is essential to have a common understanding of what data organization has and how it is managed. "The purpose of the business glossary is to drive the organization to agree on which master data is important and how it is defined, and to ensure master data is effectively communicated across an organization" [67]. It helps to link technical data components with business terminology. For example, it could be the simplest everyday thing of writing a zip code or a date in one agreed upon format or understanding what is financial year or one-day revenue across the company [5]. Business glossary creates a clear understanding of terminology, more effective use of data and "a single definition for key business terms, which fosters communication and reduces confusion about what data means" [5]. This could be a first step in building a comprehensive data architecture.

4.2.2.3 Data Asset Inventory

This is a list of all the data stores and data flows between them in both production and non-production environment. Surprisingly, there are many organizations that are not aware of the data they are storing, operating and who to contact for a certain piece of data [68]. This way it is impossible to manage the data and keep it in a good shape, the data could become redundant, trivial or even obsolete.

The data inventory needs to reflect the data structure from the smallest data store level to the high level of applications and systems. It needs to reflect what data is stored in each individual data store or disk, what kind of data is being transported between the data stores and applications. That is a complex and multi-level network, however very helpful for data management and planning for new systems or applications, it gives the ability to check for redundancy and the most effective way to get new desired data. To collect all the information

for data asset inventory system owners, users and database administrators need to be contacted. The inventory also needs to be a living document and updated frequently or whenever there's a change in system flow or data that is being gathered, especially after some major migration projects as usually some data is left that nobody knows about [5].

4.2.2.4 Data Standards

Data standards are a very important part to consider when developing a data architecture. It will set the example of the data formats, files and the way of data being transported and stored. "Detailed (or low-level) data standards might include standard terms and definitions (e.g., HL7, FIBO), standard code sets (e.g., country codes), or data exchange standards (e.g., NIEM, ANSI, ISO) [5]. Also, it could include industry-specific data standards. Data standards go beyond the raw data formats, it also includes the representation of data, so it is important to think about the standardization with the external partners of the organization that they share their data with so that it would be compatible or easy to change to a different standard.

4.2.3 Data Collection and Processing

This domain talks about the tools and procedures for data collection and processing.
When talking about data collection it is important to understand if there is a data redundancy and company is collecting the same data from different channels and within different departments [15]. Redundancy can become an issue of data quality and budget overspending. Data collection concerns the data sources and includes identification of the correct data sources and data types (structured, unstructured, semi-structured) [5]. Data sources might be internal or external, generated by the company itself or acquired from the third party. The types of data sources could include a transactional environment, reporting environment or a data lake, that is a big consolidated data source repository by itself that is becoming more and more popular [5]. The data in those data sources might be raw, cleaned or pre-consolidated, so teams need to choose carefully how their data would be collected and in what form to fit the business objectives and requirements [5].
It is quite risky and expensive to make data analysis from the raw data, it needs to be processed in a proper way to have the most value out of it [5]. Data processing includes data cleaning and data preparation for the analysis or integration in other systems. Thus, data need to be transformed for raw data into the usable data sets. The actions to reach that might include data modeling, transformation, and integration [5]. All these activities require a skilled data analyst in the company as well as selecting the right data processing approaches and tools so that they would fit business requirements.

4.2.4 Data Governance

Data governance part indicates the data owners and their involvement in data strategy, describes the main roles and responsibilities of data managers, data contributors, data approval workflows or any approval regarding the data strategy [15]. An effective data strategy is created in the collaboration with all the stakeholders in the company “to ensure buy-in and solid understanding” [69] of the strategic goals and objectives. Data governance is basically a data management part, where the defined responsible people are in control of data strategy planning, monitoring and enforcement [5]. Data governance should lead other parts of data strategy in the sense that in this part planning and monitoring should be done. In order to have a successful data strategy, the governance mechanisms need to be implemented in all data groups and assets across the company.
Data governance also addresses and impacts data quality [15]. As governance policies control and monitors data flow, data formats, data accuracy so it ensures a data quality standards in the company and also makes sure that data complies with industry standards and is able to pass data related audits [15].

All of what was said about data governance put together, it is clear that "data governance contains many aspects, which formed from a convergence of procedures, technologies, processes, policies, responsibilities, and decision-making rights for the use of data in organizations" [70].

4.2.5 Data Quality

Data quality is something that could be very subjective, however, according to the Harvard Business Press publication the quality of data is considered high when it fits its “intended uses in operations, decision making and planning” [71]. Quality of data usually resides on the expectations of the data users as well, so understanding of data quality could be very different across the organization. There are 5 data quality dimensions that would help to recognize the purpose of the data quality [5]:

- Accuracy: is related to the data value correctness in regards to the primary data source. In the organization, there could be several data sources, datasets that refer to the same value. It is important to keep data value as consistent and accurate as possible, however, data could become inaccurate within the process of transforming

and analyzing data or with the growing amount of new data sources. To prevent an organization from inaccurate data, there should be a certified or a reference data source determined to follow as an accurate [5].

- Completeness: it means that the data set fully fits the users need of having that data set. The lack of completeness would be if the data set would be missing important information and without that information data set does not bring full value [5].

- Consistency: or data integrity, it refers to the correct and consistent data representation across all the systems, datasets, applications, reports and dashboards in the organization, both internal and external. Data should be consistent not only in the value but in the format and representation as well. Inconsistent data leads to low reliability of data and of the company if it is external [5].

- Latency: this characteristic of data is related to the timeliness of the data availability. Latency indicates the waiting time that is needed from the creation of a data set and the actual representation of the data to the end user [5]. This data quality domain is very important within the companies in the financial, medical or transportation sector.

- Reasonableness or relevance: this data quality domain refers to the credibility and usefulness of the data set. The data set could have a good quality in its accuracy and integrity but without reasonableness that data set is useless. Data quality in this sense should be treated in respect to the end user, of how relevant, understandable, well documented or secure is that data [72].

Data quality requires a lot of resources, so the level of quality has to be settled in order not to apply the highest quality to all the data assets but selectively choose according to the importance and objectives of the company, otherwise this could cost too much effort and make decision makers doubt of the necessity of data quality [15].

4.2.6 Data Privacy and Security

Data privacy and security are becoming more important in the last few years, especially in the rise of privacy breaches in the big corporations like Facebook and also as General Data Protection Regulation has been enforced this year. GDPR is applicable to all the companies that are storing, processing or in any other way of operating the data [73].

As there are big fines for breaching GDPR and there is no exemption to any company, six principles of privacy should be applied to data in the organization [74]:

- Lawfulness, fairness, and transparency: the data should be processed in a lawful, fair and transparent manner, that the user should be informed about the usage of the private data. Blocking all the illegal data activities.
- Purpose limitation: the private data that is collected should only be used for the purpose that was communicated to the user and cannot be exploited in any other way.
- Data minimization: collected and processed personal data has to be relevant and adequate for the purpose of collecting and processing.
- Accuracy: an organization needs to make sure that personal data that is stored and represented is accurate and up to date. Obsolete data has to be deleted from the systems.
- Storage limitation: personal data cannot be stored longer than it is needed to fulfill its purpose. It also relates to the right of the user "to be forgotten", so the company has to be ready to delete all the personal information about the user and provide the evidence [75].
- Integrity and confidentiality: data controllers and processors need to ensure data security at all times against the possible risks and threats. Also, they need to make sure that data is not corrupted or lost [76].

Data security is a very critical part of the data strategy, which talks about the ways of protecting a company's data. Data security could also be described as getting the right data to the right recipients at the right time[5]. The widely accepted objectives of enterprise security, including data security, is 1) confidentiality 2) integrity and 3) availability [5].
For data security implementation in the company, cybersecurity frameworks could be used. According to the Mike Fleckenstein and Lorraine Fellows [5] book, most popular frameworks are NIST [77] cybersecurity framework developed by the US government and private interests or Risk Management Framework [78] (RMF) developed by the US federal government. Both frameworks are free and easily accessible.

4.3 Data as an Asset

For most of the organizations, it is natural to make money out of their assets or resources, no matter if they are physical or digital resources. However, as data is not that highly recognized as an asset, it is much harder to understand that it could also bring some financial benefit and

should be treated as a valuable resource [5]. Nevertheless, there are guidelines from the Internal Revenue Service in the US (IRS) for organizations to list some of their intangible assets as technical data or intellectual property [79]. This way enterprises might be encouraged to exploit their data in different ways.

The primary purpose of treating data as an asset is to understand the benefits and value that the data brings, i.e. increase revenue streams, increase productivity in the organization and cut the expenses where possible [80]. Treating data as an asset or monetizing it, also increases the returns of data strategy investments and convinces management to invest and contribute more to the data strategy [80]. There are many ways how to exploit data as an asset as shown in the Figure 8. [81]:

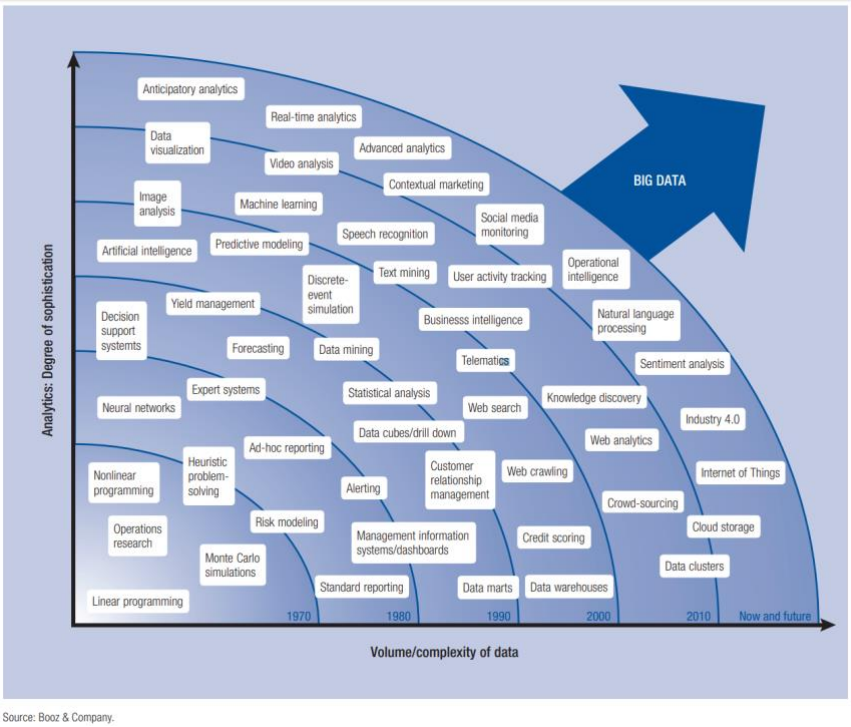

Figure 8. Evolution of data-driven decision-making [81]

According to TDWI, there are two main ways of how to get the value out of data: direct data monetization and indirect data monetization [82].

4.3.1 Direct Data Monetization

This type of data monetization approach is providing an access to the company's data for the third parties and customers. It means that the data, that the company owns and has gathered, could be sold in different ways and forms from raw data to knowledge. The efficiency of this approach could be measured very easily, as it creates the direct revenue stream. However,

selling data can cause a lot of privacy and security issues for the company, so preventing privacy breach should be seriously considered.
The ways how to monetize data directly [82]:

- Creating an access for customers to be able to use the data. Usually, it is done by creating a special API, that allows external users or third party applications to access desirable data for a certain price.
- Selling data: a company sells its data to other companies or to the data exchange platforms to generate revenue.
- Selling insights: a company can sell already processed and analyzed data to its customers or other interested parties.

4.3.2 Indirect data monetization

“Indirect data monetization requires you to think about data as a strategic asset in new ways” [82]. This approach is related to improving efficiency and reducing the cost while making data-driven decisions. Data that is gathered about the customers and sales could indicate the possibility of the new sale, upsale or the sale of a similar item. Data is processed and analyzed to help employees to make more efficient decisions, would it be related to sales or internal organizational processes. Ways of making indirect data monetization [82]:

- Discover new customers or new business models: using internal and external data, a company can discover new patterns that could become a new business model, product, service or also find out how to attract new customers to use their product.
- Enhance the product or service: discover new ways of improving the core product by analyzing customer data (customer behavior with the product and the experience or purchase) - data-driven decisions.
- Reduce cost and improve efficiency: optimize business processes according to the data analysis.

To enable data monetization a company needs to have a clear data strategy to make the use of the data they have [82] [83]. The company needs to have a clear structure of data, to have a full access for the right people, to prepare and clean data for analysis, and to be able to scale data in storage and processing.

4.4 Data Maturity Assessment

In this chapter, data maturity assessment models will be discussed as a part of the data strategy frameworks. Data maturity assessment in this research is treated as the assessment of the current state of data in the organization. Data maturity assessment is important because "companies need to understand where they stand in terms of big data maturity so that they can progress and identify the required initiatives" [84].

To begin with, a maturity model needs to be defined. Data maturity is the state of data that most of the value could be obtained [84]. A maturity model is a tool for facilitating the assessment of development level in the organization, the assessment analyses development level of processes, capabilities, and resources[85]. The same principles apply to the data maturity model, that should asses the process, capabilities, and resources of data.
Typically, data maturity models are categorized into three groups: [86]

- Descriptive: asses the current state of the data maturity
- Prescriptive: asses the current state of the data maturity and suggest the guidelines for future improvements
- Comparative: asses the current state of the data maturity and compares it with the industry trends.

Prescriptive models have been selected to investigate in this section, the ones that measure the status of the data and provide guidelines for improvement as well.
Data maturity models and assessment have not been covered in the academic papers [87], that is why the following data maturity models are developed by the consultancy companies.

4.4.1 Info-Tech Big Data Maturity Assessment Tool

This assessment tool helps to define what data is applicable to use for big data initiatives and find out what new insights could be derived from those new initiatives [88]. There are 4 phases for this maturity model:

1. Undergo big data education - this step is to understand what is big data and what it is used for, develop the knowledge of how to use it effectively in the organization.
2. Assess big data readiness - in this step the questionnaire[88] is used to assess the current state of data maturity in the organization and provide the recommendations accordingly how to move forward.

3. Pinpoint a killer big data use case - identify the case inside the organization of what data could be used to support the most important business objectives and how to use it.
4. Structure a Big Data Proof-of-Concept Project - start a big data initiative for improvements of data in the organization.

4.4.2 Hortonworks data maturity assessment

The Hortonworks data consultancy company has successfully helped over 500 customers to transform their data over the years [84]. They have also developed data maturity assessment model that which focuses on 5 domains [84]:

- Sponsorship: data strategic vision, support, and funding from top management
- Data and Analytics: evaluates the status of data collection, processing, storage, and data analytics.
- Technology and Infrastructure: addresses data hosting situation, data functionality, data integration between systems and tools.
- Organization and Skills: addresses the level of data expertise in the company, management attitude towards the data transformation and the possibility to create cross-functional practices.
- Process Management: focuses on evaluating processes regarding data governance, data privacy and security, data quality.

Hortonworks provides a free online survey with evaluating the current state and strategizing where the company wants to be in 2 years within the industry. The survey results are sent and displayed where the company stands in regards to the industry and where it is planning to go. After the survey, Hortonworks helps to develop a roadmap for the company to reach that 2-year plan or align it to industry trends.

4.4.3 SVDS (Silicon Valley Data Science) data maturity assessment

SVDS is a consulting company that provides the service of assessing the data maturity in the organization and guiding to create a roadmap for the data improvements to become more data-driven organization [89].

Before the assessment, the consulting company asks some crucial questions about data in the organization that is being evaluated, such as: "What are your organization's business priorities— and how will your data strategy support them?" [89], “Do you have the talent and capabilities required to execute your data strategy successfully?” [89], “Which areas of your

business are already data-driven? What investments are needed to support areas that are not data-driven?" [89]. After the questions have been answered, SVDS together with data engineers and scientist from the organization creates customized data maturity assessment. The maturity assessment is very detailed and touches 12 domains of data that focuses on data systems, processes and teams [89]. The six domains are shown in Figure 9.

Figure 9. Data Maturity Assesment Domains [89]

After the detailed data maturity assessment has been performed, SVDS provides comprehensive guidelines on how to improve or create data initiatives to increase data capabilities [89].

4.4.4 Ironside Data Maturity Assessment

Ironside is an experienced data consultancy company that has is providing data and data analytics maturity assessment services [90].
The consultancy company provides questionnaire of 32 questions to evaluate the maturity of data in the organization. The questions are related with 6 data domains [90]:

- Data strategy: global organizational strategy and alignment with business and IT objective
- Analytics: ability to consume company's information and create effective reporting
- Data: the status and quality of data captured, integration and transformation
- Governance: ensuring relevant data policies and standards across the whole company
- Skills: the status of data expertise in the company
- Architecture: the ability to support and scale applications and platforms

After the questionnaire has been conducted, the ability of consultation is available on how to move forward with data improvements.

4.4.5 SWOT

The well-known SWOT analysis tool could be used for the data assessment as well [58]. SWOT stands for Strengths, Weaknesses, Opportunities, and Threats that are analyzed, assessed in the organization. SWOT analysis helps to build the awareness of the strengths and weaknesses that exist in the organization as a whole entity or in a particular area, it also spots and creates a strategic plan of exploiting opportunities and eliminating threats [91].
SWOT tool works by exploring internal and external factors that could influence or interfere within the organization [91].
The internal factors are strengths and weaknesses, the external factors are opportunities and threats as shown in Figure 10. [91], that way the organization is listing their strongest skills and capabilities and identifying possible or existing issues.

Internal		External	
Strengths	Weaknesses	Opportunities	Threats

Figure 10. SWOT Analysis [91]

For the internal factors (strength and weaknesses) these areas need to be considered [91]:

- Human resources: employee skillet
- Physical resources: tools, equipment
- Financial resources: dedicated budget, income
- Activities and processes: software, procedures, operations
- Past experiences: successful/unsuccessful use cases to learn from

For the external factors (opportunity and threats) the following areas need to be considered [91]:

- Future trends: industry trends, specific trends in the field, new inventions
- Economy: status of the national or global economy, market share
- Demographics: the changes in demographics of the customers

- Legislation: new laws and regulations
- National or international phenomena

All those areas of questions should be answered by the input from the teamwork put together to brainstorm any possible positive and negative ideas.

The advantages of this model are that it is very simple, flexible and highly adaptable to any industry or managerial level in the organization [91]. It also gives a chance to see the weaknesses and threats next to the strengths and opportunities and ability to turn the negative attributes to the positive side.
The disadvantages of this model are also its flexibility and lack of description or guidelines.

4.5 Summary

In this chapter different data strategy frameworks have been investigated. It was necessary to investigate several frameworks in order to develop a strong perspective on the existing frameworks and use them later in the analysis part. Data strategy domains will also be used in the analysis part as the elements of the data strategy structure, the same as data maturity assessment tools.
Data as an asset part shows the approaches of the ways to extract the value of the data. It will be used in the analysis of the conceptual data strategy framework as the end goal of the strategy.

5. Empirical data

This chapter discusses the interviews with the case study company's managers of the different departments. There have been six interviews in total and one framework validation workshop. The first interviews with DWH manager helped to shape the idea and the problem formulation. The following interviews with different department managers helped to find out the current state of data, data strategy and the alignment with other strategies in the company.

5.1 Findings from the interview

Questions for the interviews have been formulated in regards to the literature review on data strategy framework, strategy alignment theories and similar research [32]. The main questions were: 1) *"Is there an overall data strategy in Unity Technologies?"* This question was asked to understand the status of the overall data strategy if it exists, how well know is it, who is involved in the creation of it, do people use it? 2) *"What are the biggest data challenges in the organization?"* It was important to find out the main data-related issues in the organization and how to solve them. Also, to reflect on the status of data strategy in regards with exist data issues. 3) *"Does data strategy need to be aligned with other strategies in the company?"* This question represents the focus of the research, the strategy alignment with the company.
After consolidating findings from the interviews, they are presented according to the main questions.

Data strategy:
All the respondents agree that there is no overall data strategy in the organization. However, data strategy exist as a separate element in some departments and that strategy is not broadly communicated or aligned with other department strategies. The data science (AI & ML) department shares this strategy only with the teams that have a direct impact on their operations and not the ones which have an indirect impact.
Also, there is a different understanding of what the data strategy actually is. When data science department manager was asked about this, he said that Unity has a data strategy that is called a data lake. It seems that the engineering department does not really understand what data strategy is. That it includes goals and objectives, processes and procedures, that it needs to be aligned with the business objectives not only the tool that is used for storing and processing data.

Biggest challenges with data in the organization:

The respondents were asked what are the biggest challenges with data in Unity Technologies. The answers were impacted by the specifics of the department but the respondents agreed on the main issues:

- Low quality of the data sources due to many in-house systems that are hard to scale and maintain at this level in the company;
- Immature business processes or lack of business processes - company is at the stage where releasing new products is more important than having proper processes in place. Lack of processes cost bugs and other issues, but the company is more interested in fast production, fixing bugs rather than fixing the core problems.
- Inefficient system/data architecture - there are many in-house systems that are hard to scale or integrate with other systems. As the company started being a small startup, the in-house systems suited all the needs, but with the growth of the company and sales, issues with the in-house systems grow as well. There is no coherent/central architecture that would allow fixing this type of issue.
- Lack of leadership around data - there is no engagement from top management in regards to data management across the company.
- Some of the data is ignored and forgotten. To get the best value of data, all of the data that is generated or accessible to the company should be exploited and managed.
- Unity technologies has an engineering mindset, it was started by several engineers, and this perception of building systems and not having any processes or procedures is still very viable. According to an architect, this is why Unity Technologies is successful, a company driven by engineering and innovation. On the other hand, DWH and Sales manager understand, that the company needs more effective and documented processes.

Solutions for data challenges in the company:

Most of the managers that were interviewed stated that a coherent data strategy would help to solve the majority of data challenges in the company or that data strategy would be the initial step towards finding the solution. Data Science and DWH managers also expressed their desire for the unified data platform that would be used in the company, or at least to reduce of the number of platforms that are being used at the current moment (7).

As for the data strategy and creation of data organization, that has also been proposed by DWH manager, who emphasizes on having the support and interest of the new initiative from the top management or the board of directors. Otherwise, the initiative would not even start, as the financial factor is crucial to have. To win top management over win the data strategy and data organization idea, it is needed to show what is in it for them. Meaning, to present financial numbers of cutting expenses dues to effective data strategy, or to present growing

sales due to new products created of data monetization and present what steps are needed to take to make data valuable in order to monetize it. The steps to get to the data value point would be the data strategy initiatives.
Other than the numbers, the top management could be barely interested in data strategy creation, DWH manager expressed. However, the data science manager, also mentioned that going to the new industries as medical or automotive could be the stimulus to create a coherent data strategy, as those industries are much more regulated.

Strategy alignment:
The answers are varying from no need for the alignment to the strict alignment and governance of data. However, the results vary so much as the some of the respondents were focusing only on their agenda and the objectives of their department only. But the truth lies somewhere in the middle, as most of the general managers agree that the coherent solution is needed but to the certain extent.
However, the corporate objectives and high-level business objectives are presented every quarter at the senior management meeting. Then all of the other departments have to follow those objectives, to align their own strategy with them. The organizational structure is functional and hierarchical, meaning that decisions come from the top management to the functional teams, to the department managers has to follow corporate and business objectives to create their strategy. However, corporate and business objectives are at the very high-level and the alignment is also only at the very high-level, to make the company keep growing and increase revenue by a certain figure.
The alignment between functional strategies is not happening. Some departments, have an occasional alignment between each other, but there are no processes or procedures that ensure that. Also, the alignment is happening on the conversational agreement, not documented process, so there is no assurance that strategies will be kept aligned.

The position of the need for data strategy:
So the solution of the data strategy framework needs to be very flexible and adaptable to many different profile teams without disrupting their current operations. All of the general managers stressed the point that the company is still growing at the very fast pace, so the priority is to keep growing and focusing on the revenue. A data strategy is needed and the purpose of it should be to increase the value and quality of data, so it would be able to make proper data monetization and make data-driven decisions.
Data itself is a very important factor for the company to keep growing, making the product better out of existing data, monetizing data and providing additional services to customers from the data. However, the company culture is not focused on strategies, processes, and

procedures, it is perceived more as the blockers for keeping the fast growth of the revenue. That is why the framework for data strategy needs to be so flexible and not mandatory but rather appear as the general guidelines how to treat data and the best practices in the organization in order to create value out of data.

5.2 Summary

This chapter described the main findings of the empirical data, i.e. the interviews with the senior managers. The findings are categorized according to the topics and only the relevant findings have been described. The outcome of the interview will be used to build a case study and develop a conceptual data strategy framework. All the relevant findings from the interviews will be used in the following chapters of the analysis to make the case study and develop the data strategy framework.

6. Case Study

This chapter analyses the case study of Unity Technologies company. It includes an introduction to the company, company profile, current state of data, challenges regarding data and the root causes of those challenges. This discussion will help to have a better overview of the company and to build data strategy framework for a case study.

6.1 Introduction to the Study Case company

Unity Technologies has been established in 2004. The company's main product is a platform that allows creating "2D, 3D, VR, and AR games and apps" [92]. Unity Technologies game engine is the most popular software to create video games among the developers as a third party software and has the big customers as Electronic Arts, Disney, LEGO, Coca-Cola, Microsoft, NASA, Nexon, Nickelodeon, Square Enix, Ubisoft, Obsidian, Insomniac and Warner Bros [92]. With that many big customers, Unity reaches 770 [92] million gamers across the world through the games that are made using the engine.
That is a huge amount of data that is being generated every single day. Furthermore, the volume of data that is being generated is constantly growing, especially for the past few years, as more and more game developers are using Unity engine as their primary development [92]. That is why the need for a data strategy is essential and Unity Technologies was selected as a case study for this Master Thesis.

6.2. Company profile

In order to develop a successful data strategy framework company profile needs to be investigated so that the framework would fit the needs of the organization. Findings from the interviews will be used to create a valid company profile.

Unity Technologies is the company that has been built from scratch, from the idea to the billion dollar investment. It was built on the engineering knowledge and passion for the industry. So it is rather the engineering company, that has a strong culture of innovation, experimentation, high tolerance of failure within the company. It is safe to say, that for the past 12 years the company has been pioneering: guns blazing, trying to move as fast as possible toward growing, being first in the market and in revenue. In the pioneering approach of management, there are no processes or strategies, just the knowledge, and energy to build a quality product

and service. However, as the business general manager said in the interview, the company is now at the phase where the settlers are coming in to introduce a change towards a common strategy and objectives, to clean up all the inaccuracies and misconceptions and to transform the company to more stable one, that operates according to business processes. The change moving towards the settlers approach might be very slow, as the culture of engineering is rooted in across the whole organization.

The financial capabilities of the company are not disclosed to the public, however, according to the Newzoo report [93], global games market is growing at the steady pace and the increase of 30,6 % of the global games market revenue is predicted until 2021. That being said, the company should have to have good financial capabilities and willingness to invest in the improvements and effectiveness.

The hierarchical structure of the organization is quite traditional. Vertical structure of the organization is dominating inside Unity with many different levels of management. There are at least 6 different levels of hierarchy from the executives to the student workers and even more departments and teams inside them. There are 16 different departments/organizations inside unity that could be broken down into the smaller teams. Those departments are located in 29 different offices all over the world, including different cultures and continents. Moreover, some of the departments or teams has been acquired by Unity, so there is the feeling of independence from the Unity as the whole organization in those departments, that impacts the usage of the coherent solutions as well. Having this complexity in the structure of the organization is very hard to manage or implement a global solution across all the company.

The vision and objectives of the company, according to the interviews and the press releases, are to keep growing in the revenue and market share, introduce more solutions with virtual reality and machine learning, enter new industries beyond gaming, like automotive, medicine and entertainment. To make this all happen, Unity Technologies has to have a great quality of data to be able to compete in the market of machine learning and artificial intelligence, which becomes a reason for a unified data strategy to appear.

6.3 Current state of the Company

This chapter analyses the current state of the company, what strategies are implemented, what is the relationship between strategy alignment, what are the main challenges that the company is facing in regards to data. The analysis is based on the finding from the interviews.

6.3.1 Strategy Alignment in the Company

According to the interviews that were conducted within the company and CEO interview [94] in the business journal, the main corporate strategy of unity technologies is to 1) democratize game development, meaning remove all the blocker for the developers to use the platform 2) solve the hard problems, meaning improve performance for the products made with Unity as well as going to the new industries 3) enable success for the developers [95]. All the other strategies in the company come from those 3 main goals.
There are yearly and quarterly meetings with all of the heads of the departments to represent their goals and objectives for the upcoming year and update on the status in the quarterly meeting. During those meetings, the department roadmap gets approved or amended by the company's board of directors. Also, that is how the strategies get aligned with the other department strategies, as business goals are presented and the input from IT goals and objectives are taken into account, and vice versa.
However, there is no overall data strategy in the organization. Data strategy exist as separate parts in the different departments and it is not globally shared across the organization. The positive side is that every department has a data strategy that fits only their needs and they can focus more on their own objectives. The negative side is that the data strategies are not aligned in the company and it can create a lot of data challenges that have been mentioned in the problem area chapter of this thesis.

6.3.2 Data Challenges in the Company

This chapter discussed the challenges that Unity Technologies has related to data and strategies. The challenges were collected from the interview with DWH manager. So what are the biggest challenges that Unity is facing regarding data?

6.3.2.1 Quality of source data

According to the DWH manager, there are plenty of source systems in the company that generates data. The issue is not with the number of source systems, but the fact that some of those systems generate the same data, but the value of data sometimes could be different. Then it takes more time to figure out the correct value and fix it. Also, the data come with many different formats, there is no standardization for the data formats, so DWH team has to integrate these different types of format and create valuable reports, which takes longer time and more resources.

The other part of this challenges comes from the implementation of the source systems. Unity was a small company and had a capability to build their own in-house systems that would accompany the company's needs at that time. However, when the company started to grow and especially in the speed that it became a company with 2.7 billion users, all the in-house systems had to scale. However, it is very hard to scale the in-house build systems and integrate them with other systems. The systems were built on top of other each other, which caused a chaotic implementation. This kind of implementation causes some errors or creates troubles with data integration and data quality, the issues are not critical, it just takes more time and resources to fix them.
At the moment there is still some internally built application, however, the company is going to implement new applications that are trending in the industry and that has a good engineering support to ensure the quality of the source systems and the raw data.

6.3.2.2 No coherent data strategy

DWH manager, as well as other managers that there interviewed, talked about the lack of a unified data strategy. At the time being some departments have some pieces of data strategy that is not shared with the other departments, it's entirely within the specific team. And some other teams even do not have any strategy to follow. This creates a lot of issues with data integration and even access to the data. By not knowing what vision other departments have on the data, it is hard to even know what data are they operating, what platforms or tools they are using. It creates a lot of duplicates and unnecessary work. As DWH manager expressed, Unity now has 7 or 8 major data platforms that some of them are doing the same operations. Platforms that are build reflect the same data, so there are several copies of the data, which becomes a problem with aligning and selecting the correct final data for processing it further. Basically, each data related department is buying or building its own data platforms to serve their needs. And in most of the cases, one unified platform could serve the needs of several teams. It is even more significant with data analytic tools. Each department is buying a different tool and paying for the license instead of using the same one. For example, the company is using several applications for data visualization and analytics. So there is no standardization in the company in terms of which software need to be used. It is not a negative thing, it creates a freedom, fulfilment of specific needs and independence in the decision making for each. However, it creates a financial burden and some issues with data compatibility between systems.

As there is no unified data strategy, there is no coherent understanding of the quality of data. Meaning that each team interprets the quality differently. Usually, quality is measured by KPIs which could be included in the overall data strategy. By not having a coherent understanding of how the data should look like, some teams could have challenges by interpreting another's team data.

6.2.2.3 Immature business processes

As the company is still in the maturing and it is somewhere in the middle of being a startup and becoming a growth/establishment company [25]. Some new departments or business operations are still being established, so the company is relatively young. So the business processes are still there new or sometimes there are no unified processes, as it was mentioned before, each department has its own way of performing operational work. Very important part of the process for data is the requirement gathering and if it is not clear and understandable for all of the teams it creates some data quality issues later.

6.3.2.4 Data architecture

As it was mentioned in the previous chapter about data roadmap, data architecture is one of the main domains of the data strategy as it sets the data flows and relationships between the systems. However, there is no overall data architecture in the organization. Every team or region has its own data architecture which is not necessarily coherent or compliant with the data architecture of others teams, which means there is no integration within the systems or a broader map of all the system landscape.
The business glossary is one part of the data architecture, which defines what data and how it is managed in the company, provides the unified terminology. In the analyzed company, there is no unified business glossary, which creates some challenges. For example, the definition of "active seat" in different departments means a different status, it could create miscommunication or even misrepresentation of factual data. Also, there is no data standard in the company of data formats, that creates extra effort in combining several formats into one coherent system.

6.3.3 Root causes of the data challenges

To find the solution for the data challenges and implement a new data strategy, a company needs to understand the main causes of those issues and challenges. One of the questions

in the interviews was regarding the source of the problem, how did it happen that there are challenges at the current state of the company.

The first and most important cause is that there is no department or team that would be responsible for the overall data issues, maintenance, and improvement. That team would be called data organization, as it was referred like this in previous chapters. Lack of the responsible team across the company creates misalignment and inability to see a big picture and value of the data.

Lack of top management attention to the data management/governance projects. As written before, the focus is on the growth of the revenue and not of the new procedures. However, if the data strategy framework is presented as effective in decision making and money saving, the focus and priority could move to the data processes and procedures implementation.

In-house built systems are also causing a lot of data challenges in the organization. According to the interviews, as Unity Technologies was a small company, it was easier to build their own systems and maintain them. However, it is much harder to scale and integrate with other systems.
All these root causes should be addressed in the data strategy to find a solution in order to solve most important challenges with data in the company so that the company could successfully implement new improvements and solutions related to data.

6.4 Summary

In this chapter, the study case has been introduced. The introduction and description of the company, company profile, data challenges and root causes of it, the current status of the strategy alignment in the company have been discussed. Some of the information has been taken from the literature review but the majority of the information from the interviews.
The case study has been built to apply it later in the conceptual framework analysis chapter.

7. Conceptual Framework Analysis

In this chapter the conceptual framework of the research will be discussed, how to construct it and apply for the case study. A lot of companies that have tried to implement big data projects or programs has failed due to the lack of effective data implementation framework [96]. That is why in this chapter the conceptual framework will be analyzed, of how different theories, methods, models, and approaches could be used to combine into one final data strategy framework. For this particular framework strategy alignment theories, data strategy frameworks and data strategy domains will be considered to make up the final conceptual framework to fit the case study of Unity Technologies.
Graphical representation of the framework will also be presented and explained in details as the guidelines for data strategy framework.

7.1 Strategy Alignment

The main strategy alignment theories have been discussed in the previous chapter. This part will discuss and select the best applicable pieces from the theories to the case study.

7.1.1 The Three Levels of Strategy

This theory talks about how the strategic goals should be alignment through all the levels in the organization: corporate strategy, business strategy, and functional strategy.

As Unity Technologies has a hierarchical organizational structure, the theory of three level of strategy could be applied. The corporate objectives of Unity Technology are the highest level of objectives, more like the mission and vision of the company, and it is implied in the business objectives that every other strategy and objectives should follow. The data strategy and other department strategies lie under the functional strategy level, as they are much more detailed and have a specific scope and function.
All the functional strategies including data strategy should support and be aligned with business and corporate objectives. The concept of strategy level alignment will be adopted in the final conceptual data strategy framework as it follows the levels hierarchically from the top: Mission/Vision (corporate objectives) is the highest level of objectives that lead the way for business objectives, and business objectives lead the way to the data strategy objectives. This way all the objectives will be aligned across the strategies.

7.1.2 SAM Model

Strategic alignment model (SAM) also talks about aligning business strategy with other strategies in the company, the same as three level strategy. However, the SAM model additionally talks about the interconnections between different strategies and their domains in the company. That would mean the alignment between the functional strategies as well as business and corporate.
For creating a data strategy in Unity Technologies, the alignment of the strategy's objectives should happen with each major department strategy to make sure that data strategy would have a coherent solution across all organization. The integration between different domains, processes, operations, and data strategy would have to take place as well, for example, IT strategy domains should be aligned with data strategy domains, such as scope and capabilities in order not to have issues with data scaling and quality of source data. The alignment between different strategy domains and the alignment with domains within the strategy will be illustrated in the conceptual framework.

7.1.3 Co-evolutionary IS/Business alignment Model

Co-evolutionary model, as well as two previous models, have alignment between the strategies and operational domains. However, it also talks about the alignment between systems and end-user needs and how they are coevolving with each other. This model touches all the levels in the company from the highest strategy to the end user. For Unity Technologies it could also be applied, the better user and system interrelations between one another, the fewer issues experienced with data quality. Processes could be built more efficiently if it's aligned with the user needs. The user also is less likely to make mistakes when they are introduced to the correct process.
In the conceptual framework, the individual level of Co-evolutionary IS/Business alignment Model will be addressed as the data and the end user alignment to gain the value of the data on the individual level.

7.2 Data Strategy Frameworks application to the case study

7.2.1 ThotWave’s Data Strategy Framework

This framework is flexible and simple enough to be applied to the case study. The first steps of the framework could be used to apply to the company 1) develop the strategy 2) create data strategy roadmap 3) align with other strategies in the organization. They are quite general to fit any company and also uses theories of strategic alignment to make strategy coherent within the company. However, the steps of implementation of strategy and monitoring, adopting the strategy should be skipped for the case study. This is due to the reasons that the company is focused on revenue growth and not on implementing strategies and procedures. Therefore, the steps for data strategy should be the guidelines or recommendations to the data strategy adoption to the company.

7.2.2. Three Tiers Data Strategy

This strategy encompasses the vision of the company for the data strategy, it focused on the data as the value and tries to get most of the data according to the company profile. The framework is using tiers for the different approaches of data strategy. For this study case of Unity Technologies, the mix of all the tiers would be the most applicable approach. Unity can use the approach and vision from all the tiers as it seeks using data for better decision making (tier 1), re-building or improving data infrastructure (tier 2) and creating new business models out of data (tier 3). From this framework, the following parts could be adopted:
Tier 1: treating data as an asset or a capability rather than a resource and improving data analytics that would help to make decisions that are more effective and increase sales. This concept fits the end goal of the framework of making data more valuable. Tier 2: only the small part of the focus on infrastructure could be adopted as Unity Technologies is not in the industry on infrastructure, however, the company owns some and manages all IT infrastructure. The improvements in the infrastructure have to be addressed to make the service more satisfying for the end user and improving the data quality and integrity and well. However, the technological solutions are the hot topic in the company, and should only be kept on the suggestion level. Tier 3: for Unity Technologies the data is not the center of their strategy as this tier suggest, but there are ideas to monetize data in more ways and make new business models by using existing and introducing new data flows. This idea also contributes to the end goal of the data strategy, to make data more valuable.

This Three Tiers framework is illustrating only one part of the data strategy, vision, which indicates where the company is able to be or wants to go. However, there is no actual action how to get to that vision in this framework, so it could only be used as a starting point to find out company's vision of data strategy, which in this case is the data value.

7.2.3 Harvard Data Strategy Framework

This framework talks about two approaches how to treat data: data defense or data offense. Defense is focusing on data control and governance and offense are focusing on data improvements to increase sales and profitability. As the framework suggests, there should never be a 100% of any approach, but in the case of Unity Technologies, the offense strategy mechanisms are much more applicable due to the company profile to keep the growth. The defense mechanisms should be applied as well, just to the extent of GDPR compliance [97], so the privacy and security data strategy domains. This domain should be based on the “single source of truth” (SSOV), especially as the company has addressed the challenge of quality and integrity of data sources. As the framework suggests, the primary sources of data should be SSOV and unchangeable to keep consistency between systems, so the defense mechanisms need to be applied to the sources of data.
Offense approach supports “multiple versions of truth” (MVOT), which ensures flexibility for the engineers and developers. In that sense, the solution for the data platform should be flexible and easily adaptable for everyone.

7.2.4 Eckerson Ten-Step Process

This framework seems to be the most appealing from all of the frameworks that have been evaluated before in this research for the applicate on the study case. The framework is very agile and straightforward so it would be easier to adopt in the company like Unity Technologies where the focus is not on creating procedures but rather to have the recommendations or best practices that the team could follow. Eckerson framework has some additional steps that other frameworks have not even considered. Building awareness, educating the team, assessing the current state, preparing for the change - these steps are essential for a company like Unity Technologies, as it allows creating a starting point for the culture of letting in the processes and procedures.
The framework suggests developing a business case where a detailed plan for executing data strategy would be created and approved by the top management. In the study case, as advised in the interviews, it is better go be even more agile and adopt the strategy on the go and not to have a very detailed action plan. As the company is so complex and there is no team responsible for data management, it is better to start small and then to evolve with the feedback from the stakeholders.

7.2.5 SAVV-BIGD Framework

This framework has been developed for the big data strategy implementation for the purpose to visualize the value of data in the organization. This framework already combines the strategies for the alignment and big data implementation steps that have been discussed in the previous frameworks, thus could be adaptable for the case study of Unity Technologies.
As other frameworks as well SAVV-BIGD emphasizes on the big data strategy alignment with business strategy alignment. So this framework already has all the parts as previous frameworks, as it is using the same theories that have been discussed in this research as well. However, the main purpose of SAVV-BIGG Framework is to clarify and illustrate the value of data for the company. It verifies the main objective of the data strategy, to increase data value and shows the relevance and importance of the strategy alignment for the data initiatives.

7.3 Data Strategy Graphical Framework

After analyzing data strategy frameworks and strategy alignment theories in regards to the case study, a conceptual framework has been developed. Below is the graphical representation of the conceptual framework (Figure 11). This framework is the recommendation for the company on how to build a comprehensive data strategy across the company. This framework represents the steps that are needed to take in order to reach the end goal, to increase data value for the company. The end goal of the framework was set after the interviews and framework validation workshop. The end goal should be related to the value of the data itself due to the top management validation of the data strategy.
A graphical framework should not consist of heavy technical terminology and should use as little words as possible so that the idea of data strategy would be easily understandable for all the departments that include non-technical people as well [10].

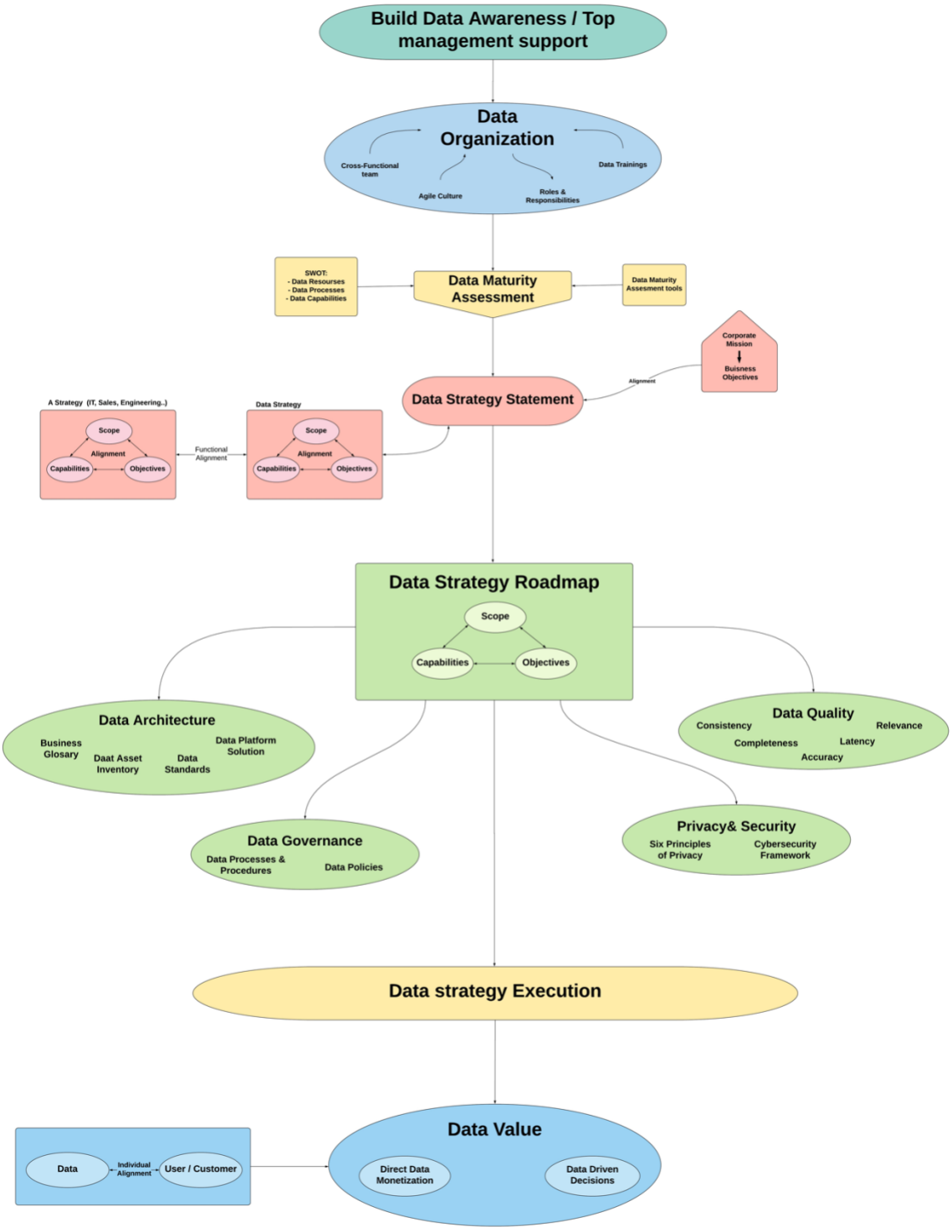

Figure 11. Data Strategy Framework

7.4 Data Strategy Framework Guidelines

In this subchapter, steps of the graphical framework will be explained in details as well as the applicability for the case study.

7.4.1 Build Awareness / Top Management Support

The first step, which is illustrated in the Figure 12., starting any new initiative is to introduce all the people in the organization, especially the decision makers, that there is an issue in the organization and a possible solution to fix it. Building an awareness of the problem and the solutions could be started with talks in the company's monthly meetings or special events about new initiatives that Unity Technologies are having frequently. This step is a chance to show different departments how they can benefit from the right way of managing their data without being too persuasive in order not to receive a backlash because Unity Technologies is the company of engineers who are quite hard to engage in new strategies, processes or procedures.

However, it is a chance to develop an interest especially if the example of a working prototype is presented of how the data could be handled in a more effective way or to show how other companies are using data for market advantages. If this step is skipped, it might be very hard to implement a new strategy as the majority of people would hear it for the first time and it would be just another procedure that would not seem as relevant for them.

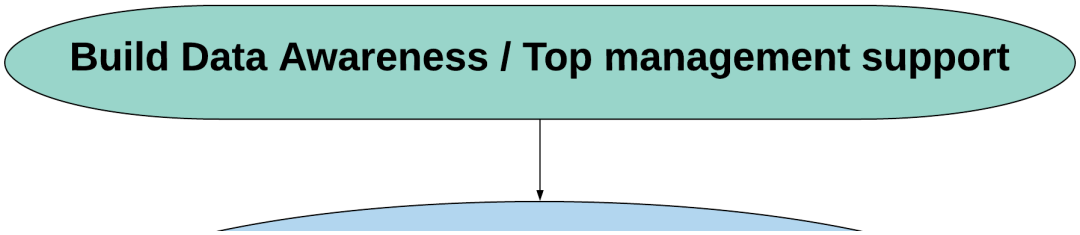

Figure 12. Data Strategy Initiation

Not only it is important to get the organization aware of the potential data strategy but to get the recognition and support from the top management and to have the approval of starting the initiative and fund it. According to the interviews with senior managers, to gain the support and financing for the initiative, it is important to show the value of the project. Top management is interested in the added value only, so to gain the support from them the whole data strategy plan has to be presented in terms of how the value will be added. The end goal of the strategy should be clearly stated, that a data strategy would increase the value of data to be able to monetize the data and make data-driven decisions.

However, in the case of Unity technologies, top management could be also interest in creating a data strategy and data organization if communicated in the right way. If Unity Technologies wants to go to new industries, such as automotive or medical that are strictly regulated, the

company needs to become more process oriented and standardized. Then building a data strategy becomes a necessity in order to have coherent data processes and procedures across the company.

7.4.2 Data Organization

Once the awareness for the need of data strategy has been built and the support from the top management has been received, a team responsible for creating a data strategy can be assembled. This team can be called data organization, as it would be not the department but the organization is assembled for one purpose: to create data strategy vision and objectives, a plan for data initiatives and for execution and monitoring. In the Figure 13., it is illustrated what should be considered for data organization.

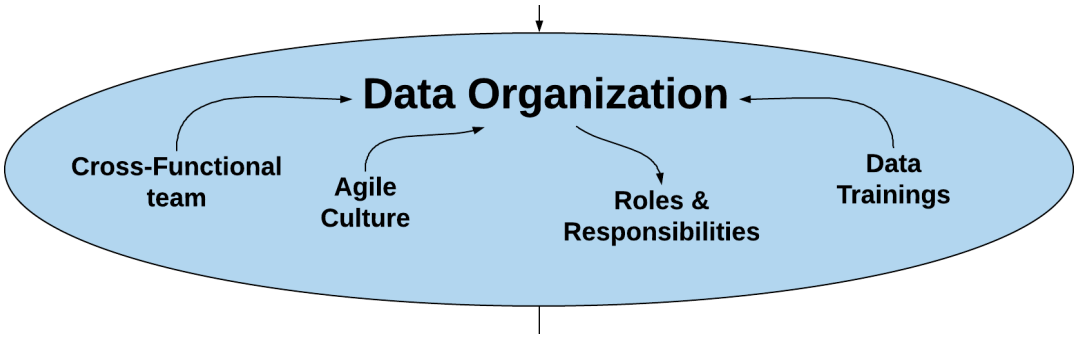

Figure 13. Data Organization

To form a Data Organization it needs input: cross-functional team (stakeholders from different departments), working in Agile culture and various Data Training. The output of the formed team is their roles and responsibilities.

A cross-functional team is needed for creating a global data strategy, that would include all the stakeholders from different departments. It might not be easy to assemble this team in Unity Technologies as it is located across many different countries, however, the idea of the virtual cross-functional team could be considered. At least this way all the opinions and needs in regards to data could be addressed.

The team should work in an **agile manner**, as it is highly acceptable in the company, however, they should agree on which agile method to follow as different departments use different agile methods across the company. Nevertheless, it is recommended to use Agile Scrum [98] method for the cross-functional team, it is a comfortable method that encourages to build, develop, test and implement a working item when it is acceptable [99]. Also, several teams that have been interviewed are working in the scrum sprints. Working in agile scrum methodology means, working in iterations/sprints (could be from 2 to 4 weeks) and after each

iteration to have an actual deliverable [99]. As illustrated in the delivery process Figure 14., the autonomous cross-functional team would gather together to generate ideas for the data strategy and create a backlog of work actions to achieve the objectives. The team would work in sprints of 2 to 4 weeks long and would have a daily stand-up for the status of the ongoing sprint [98]. The frequencies of the sprints and status meetings could be adjusted according to the capabilities of the team. Using the Agile Scrum methodology, the team can reflect the delivered pieces from the stakeholders' feedback and from monitoring the already implemented solutions to improve them. Scrum methodology should be used until the team has any working items, just the frequency of the sprints and meetings might chance.

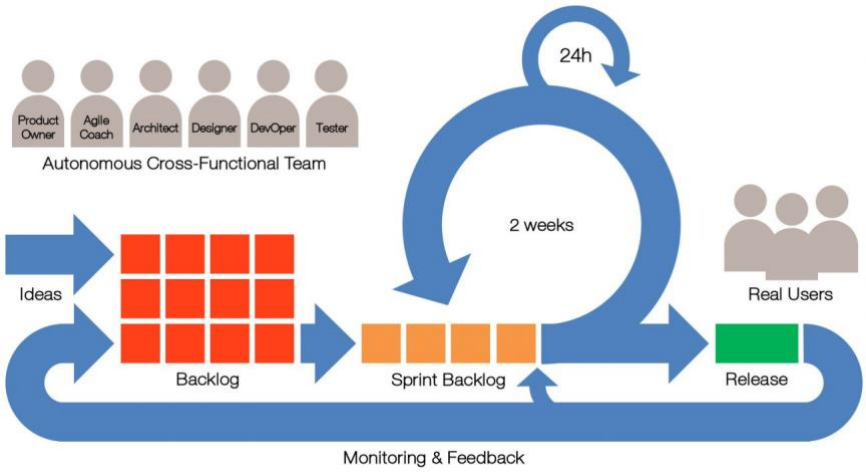

Figure 14. Scrum Agile Process [100]

In the agile cross-functional teams also means that each team is responsible for the different feature of the deliverable. In this case study it can work the same, the security manager and IT manager would have different roles and responsibilities in the team, where one develops requirements for data privacy and security and the other of the new or improved infrastructure. The **roles and responsibilities** need to be clearly described and documented of each stakeholder of the data organization. However, the cross-functional team would still do the planning and strategy development together. It is also important to select the right people to become members of the data organization. It should be the representatives or the senior managers of each department that are working with data directly in the organization [58]. There might be an issue that some of the departments would refuse to participate, but if there is a strong top management support it should be easier to get all of the departments on board.

When the new data organization has been established, it needs to be **data trained** with all the information about the data and its strategy. The company should hire an external consultant to help with the training of the team. As in Unity Technologies case the team is global and scattered across the continents, it should be gathered in one physical place for the introduction

and training workshops. Later on, the team could work virtually but at least for training, they should gather in one place.

7.4.3 Data Maturity Assessment

In this part, the cross-functional team should assess the data maturity in the company as shown in Figure 15. As it was listed in the research chapter, there are some tools for assessing data that could be used, external or internal depending on the financial capabilities and approved budget.

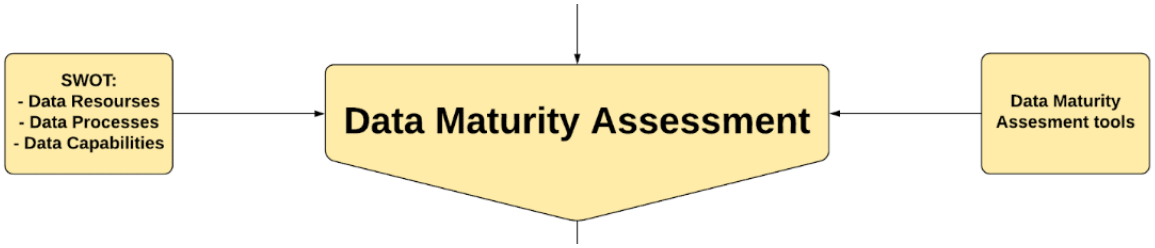

Figure 15. Data Maturity Assessment

There are advantages and disadvantages of using external assessment tools from consultancy companies that have been discussed in the SOTA chapter. The advantage is that managers of the department do not need to spend too much time thinking about the assessment, but just fill in the questionnaire and wait for recommendations provided by the consultancy company. Also, those companies are professionals in the field, so they would be able to guide the required action with the expertise. Furthermore, those tools touch all the necessary areas of data: status of data architecture, data processes, data privacy and security, and the data quality. However, the downside of this type of assessment is that it provides a general type of questions that might not be suitable for the specific department in the company and also the tools come with additional fees, so it requires budget spending.

SWOT analysis is the internal option for an assessment a tool. SWOT is very flexible and adaptable to use, that allows creating a picture of positive and negative forces in the organization addressing the issues that are specific to each team. While using SWOT the focus should be on evaluating data resources, data processes, and data capabilities according to the principles of maturity models [85]. Conducting a SWOT analysis takes the following steps, that have been created according to the general SWOT analysis guidelines [91] and applicability to the case study:

1. Brainstorming all the positive and negative forces in the organization
2. Review the list with the team and leave the most relevant points

3. Link the negative aspects of threats and weaknesses to the positive strengths and opportunities to create a possible solution
4. Create an action plan for exploring the opportunities and mitigating the threats
5. Assign responsible people for the action items
6. Follow up on the plan

As advised in the interviews and framework validation workshop, for the most successful result, data organization should use both, external and internal tools of data maturity assessment. However, the primary data assessment method should be the professional assessment tools and the SWOT analysis should be a supportive method for the assessment. In this way, the company would have the view and recommendations from the professionals and an internal judgment from data experts within the company that knows all the internal data issues and strengths the best. It might take time to asses all the data in the company but this step is crucial because as the famous quote goes: "what cannot be measured, cannot be managed" [101].

7.4.4 Develop a strategy statement

After the data maturity assessment has been conducted and action plan developed, the data strategy statement could be created. However, before developing the statement, data organization needs to look back to the corporate mission and company business objectives. As all of the discussed theories says, data strategy needs to be aligned with the mission of the company and its business objectives, that the company would be able to go to the direction that is agreed upon.

In the strategy statement, there should be data strategy objectives, data scope and data capabilities outlined as it is suggested by almost all investigated frameworks and theories.

- Data strategy objectives that represents the main data strategy purposes, and the end goal of the strategy, data value :“to make data-driven organization” or “to monetize data in the organization”, these objectives will lead data strategy for the next upcoming years until the objectives are reached or changed [15].
- Data scope includes the boundaries of data strategy, which areas does it affect or what kind of data will be included.
- Data capabilities represent the competitive advantage or uniqueness of the organization in the market [15].

The elements of the statement should also be aligned with each other, so the strategy would be consistent. Data strategy statement could be realized into a written form of 3 combined

data statements: scope, capabilities, and objectives. Part of that and the whole strategy statement formation is illustrated in Figure 16.

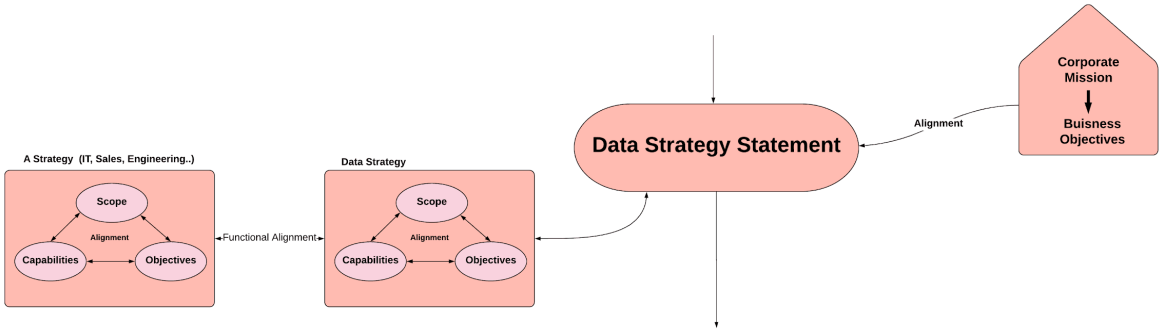

Figure 16. Data Strategy Statement

Functional alignment

After the data strategy statement has been established, the alignment with other strategies in the company is needed. For example, the data strategy statement should be in alignment with the IT, Sales Operations, Engineering departments strategy statements. The alignment also goes back to the data strategy elements. This approach is suggested in the SAM model for a coherent data strategy across the company. Strategy statement functional alignment is also the reason to have cross-functional teams in the data organization then is much easier to alignment all the strategies to reach common objectives.

Scope

The scope of the data strategy is an important thing that needs additional attention and that is created in this step of the framework. It should be one of the first data strategy elements to consider. The importance to start with the scope is the allocation of limited resources, so the organization needs to categorize their focus of data. As DWH and data science department managers said, all of the data should be included in the scope but with different priorities.

According to the Thotwave guidelines, data should be categorized into two main scopes: operational data and data for the improvements. However, as Unity Technologies wants to create a customer 360, those two categories overlap, so it would be better to categorize data according to its purpose and priority.

7.4.5 Develop data strategy roadmap

As it was discussed in the data strategy domains, there are several key elements that need to be considered: data processes, technology, and people. For those key elements, there should

come out different data initiatives, that make up the data strategy roadmap. The roadmap or the action plan for Unity Technologies could be done from several data initiatives that are indicated below. However, it is only the initiatives that are recommended and seems relevant for Unity Technologies after having all the interviews with different department managers.
The scope, capabilities, and objectives are also in the alignment, as shown in the Figure 17., and also it was already mentioned in the data strategy statement.

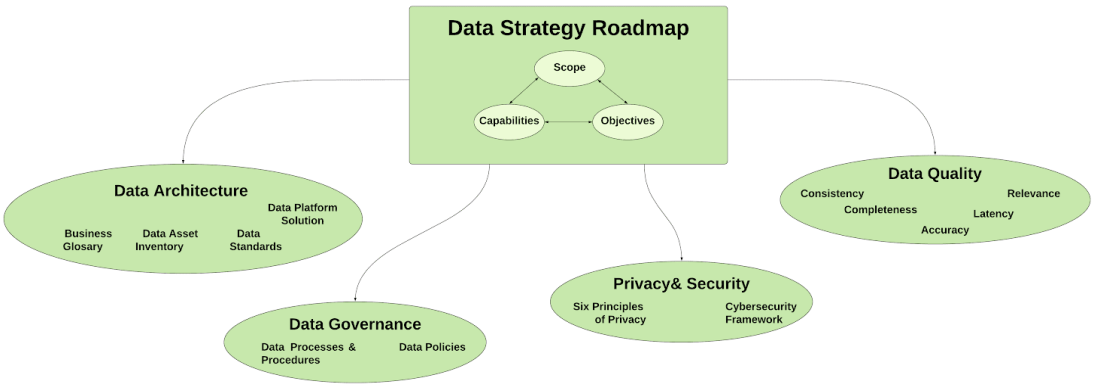

Figure 17. Data Strategy Roadmap

7.4.5.1 Data architecture

Data architecture should be restructured as it causes many challenges with data integration and data quality. Unity Technologies does not have the overall data architecture in the company, which creates a gap for issues and challenges to happen. The coherent solution would prevent a lot of problem for later data processing.

- A business glossary should be established to address the common understanding of the terminology across the company, as in the interviews it was found that different departments have different definitions for the same phrase. Business glossary needs to be created with the input and understanding from all the stakeholders.

- To create a data asset inventory should be a second step for Unity Technologies in data architecture. This is the list of all the data stores, data flows, data tools, data applications and data procedures that exist in the organization. According to the interviews, this type of list does not exist in the company at the moment and it causes redundancy of the systems and resources for the same purpose. If there is no visibility of existing data capabilities, then each department would start to build their own

solution, which would end up in budget overspending and possibly other data integration issues. So this step is also crucial for the case study company to have.

- Data standards should be considered as well. It will save time and money for later data processing activities such as data transformation and integration. However, as the solution architected stated in his interview, standardization of data would not be a good way to go for Unity Technologies. The freedom of choice for each team is what makes Unity Technologies a unique place and helps to create a unique product. Data standardization, according to the architect, would restrict and limit teams from creativity and innovation which are the biggest values of this organization. Thus, data standards might be considered but should not come as an obligation.

- Finally, a coherent data platform solution. Coherent means, that there are several data platforms with the clear understanding when and how to use it. Data lake should be the primary data repository and data processing platform for raw and unstructured data, except transactional data due to the security reasons. As the manager of the Data Science department expressed, Data Lake is the strategy that the company is following, however, as it is not broadly communicated to all the departments, so they are using different data platforms or even building their own data platforms. Data lake could be used both for data collection and data processing. The transactional data should go to the EDWH platform where it would also be processed and analysed.

7.4.5.2 Data Governance

Data governance refers to the planning, enforcing and then monitoring all the data management areas. Meaning, that it takes part in the data strategy planning, enforcing an execution and monitoring the strategy. To that sense, data governance should be embedded in all the part of the data strategy, however, data governance also needs to stand as a separate data initiative because all the processes, procedures and policies regarding data should be created and monitored within this initiative.
For the case study of Unity Technologies, data processes and procedures are not the first priority as it was mentioned in the interview part. So the processes and procedures that would be created should not be enforced but rather recommended as the best practices in the company.

However, there are some procedures that cannot be skipped due to laws and regulations, for example, GDPR forced to have data security and privacy procedures that are mandatory, in order to comply with the regulation.

7.4.5.3 Data Quality

Data quality initiative addresses the issues identified in the data maturity assessment. From the assessment, there should also come the action plan to improve the quality of the data. As it was discussed in the literature review, data quality initiative should focus on the 5 elements: Accuracy, Completeness, Consistency, Latency, and Relevance. So the action plan should be adjusted to those 5 elements of the data quality, as well as the scope, capabilities, and objectives of the whole data strategy.

7.4.5.4 Data Privacy and Security

Data privacy and security should be handled by the security department. Because of GDPR, a new role of risk and security compliance manager has been established to prepare the company for GDPR. They have been working to be compliant with the regulation for the past year to make sure that there are no security or privacy breach. Nevertheless, privacy and security are very important elements, so have to be included in the data strategy roadmap.
The assessment of privacy and security might be done in the data maturity assessment, if not then there should be a separate assessment and then the action plan in regards with six principles of privacy from GDPR and cybersecurity framework that have been described in the literature review chapter.

7.4.6 Data Strategy Execution

In the literature review, data strategy execution is explained as the implementation of the data strategy in everyday business and operational workflow. The plan/process for data strategy execution should also be included in the data governance initiative.
However, as it was mentioned in the interviews and case study analysis, the enforcement of the strategy implementation is not in the culture of the company. In this case, data strategy should be communicated as the best practice example by the stakeholders of the cross-functional team (each stakeholder is the manager of the department in the company). And the initiatives that could be started, should be implemented into each department respectively.

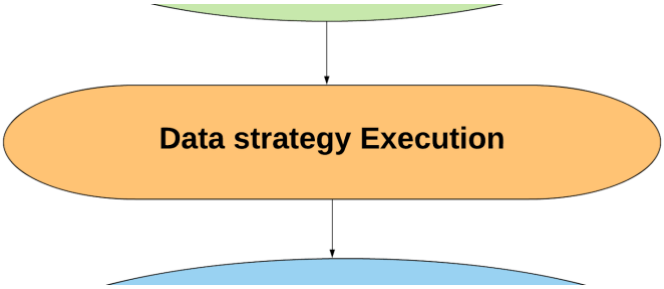

Figure 18. Data Strategy Execution

7.4.7 Data Value

The data value is the endpoint or the goal of the data strategy for Unity Technologies. All the steps and action plans that have been described above and potentially implemented, should lead to the purpose of making data more valuable to the company and treating data as an asset. That is the main data strategy objective due to the culture of fast-growing companies, there is nothing more important than making more revenue. As the DWH manager expressed, the purpose of being compliant, aligned and more efficient with data would not be the reasons that top management would support, especially financially. The only way to go is focusing on the data value and how it could increase revenue.

According to the literature review, data as an asset comes in two forms: direct monetization and data-driven decision making (indirect monetization). There should be a data customer monetization department which could manage customer 360 profile creation initiative, that could be sold to gaming companies, it is only in the idea level and there is no responsible team for implementing it. Indirect data monetization or data-driven decisions could be a part of the data science team, as their manager suggested and because they are building new machine learning technologies that could improve internal processes for different departments. All of those improvements should come from aligning end user or customer needs with data.

However, to actually implement data monetization and data-driven decisions all of the data strategy steps are needed. Data strategy would solve the majority of existing data issues in the company, ensure the cross-department alignment with the data needs, and makes data more accessible and better quality to use.

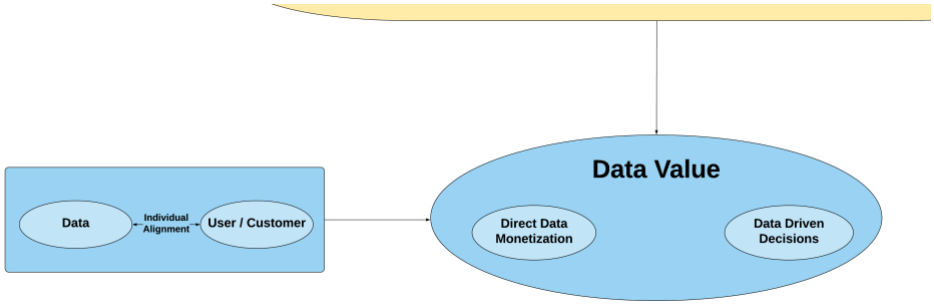

Figure 19. Data Value

7.5. Data strategy framework validation

The framework has been validated with DWH manager who was the first initiator of the data strategy in the company. There was a workshop in which data strategy framework has been presented step by step. In the beginning, the framework had more technical and detailed solutions for data architecture, however, as DWH manager explained, Unity Technologies is the company of engineers, so to suggest a technical solution could be very tricky and there could start big discussions and even resistance against the framework. So the technical architecture part should be excluded from the data strategy framework.

Also, DWH manager has stressed a lot on the essential support from top management and that they have to be pursued with numbers or something really crucial for the business so it cannot be rejected. Otherwise, just with the purpose of making data more effective, does not sound too convincing for the management to put money and priority on it.

The last suggestion from the DWH manager was to put data value in a separate graph, which it would clearly show, data value as the final goal of the strategy. It also relates to the part of having the support from the top management. This way the reason to have a data strategy comes to the purpose of making data more valuable for monetizing it and creating a data-driven decision.

The rest of the framework was accepted by the DWH manager. He stated that the framework could be a good start point for implementing a data-driven culture in the organization.

7.6 Summary

In the Conceptual Framework Analysis chapter, the data strategy framework has been constructed and explained in details. The construction on the conceptual framework started with the analysis of strategy alignment theories applicability for the case study. Parts of the concepts have been taken from each theory to construct the final data strategy framework. After the analysis of the theories, existing data strategy frameworks have been analyzed for

the application to the case study. It was explained which parts of the framework are applicable and why. The applicability of theories and existing frameworks for the case study has been measured and analyzed against the empirical data and the company profile.

After that, a conceptual framework has been constructed for the single study case and presented in a graphical way. The framework represents all the steps that are needed to take to reach the end goal of the data strategy framework, to increase the value of the data in the company. The conceptual framework consists of eight stages which have been described in detail and with an explanation of the suitability for the case study.

Finally, the internal validation workshop of the data strategy framework is presented. Validation has been performed with the data warehouse manager, the same manager that has introduced the data issues in the company and inspired the problem formulation. During the workshop, changes have been made for the framework regarding the initiation of the project, data architecture and the end goal of the strategy. The conceptual framework has been adjusted according to those changes and framework accepted as it could be the starting point of change in the organization towards data.

8. Discussion and Conclusion

8.1 Discussion

Data has become an important part of every organization life. It is not enough just to use the data for reporting a company's finances or sales. Data is being used in more sophisticated ways and organizations are even selling the data or making data-driven decisions. In order to reach that level of data management, an organization needs to develop a data strategy. However, only a very small percentage of the companies are using their data in the way that it would become valuable and beneficial. So this research has a high relevance among the companies that are using or generating huge amounts of data.

The main objective or the end goal of the research was to develop a data strategy that would be suitable for all the departments in the organization. To build that data strategy, a case study has been used, so that the strategy would be more applicable and relative to the data issues that exist in that company. A single case study for Unity Technologies has been used for investigating and constructing the framework. By using a single case study approach, the research had more depth and insights from the company with the actual issues. The case of the company is not very common, however, it is related to the fast-growing companies. Because of that, the existing data strategy frameworks were not sufficient enough to fulfill the needs of the case study and a conceptual framework method has been chosen to build a data strategy framework. The conceptual framework is highly flexible and adaptable by using combined theories and models. Conceptual data strategy framework has been constructed in regards to the discussed theories and literature review, as well as empirical data from the interviews with the case study company. It helped to build the framework that is customized and that fills the data gaps in the company by offering alignment of the scope, objectives, and capabilities between different departments, as well as the alignment with business objectives to ensure the data initiatives are going along with the company's mission.

The constructed data strategy framework could become as the initial step towards changing the culture of data management in the organization. The graphical framework could be presented to the top management for their support, as it shows the concrete outcome of the strategy: data monetization and data-driven decisions. Those two outcomes are the indicators of the new and increased revenue, monetizing data by creating new products and data-driven decisions assist in better and quicker sales. This is the value that brings revenue and keeps the company growing.

Implementation of the data strategy and its initiatives should not be obligatory for each department. Unity Technologies is the company of engineers and is used to build and operate systems without any processes or procedures. Procedures are seen as the blockers for innovation and fast growing. However, Unity Technologies is maturing and leaning towards having effective business processes, so the culture is slowly changing but not enough so something that would disrupt. That is why the data strategy framework implementation is not discussed in this research, it should be left for individual departments to decide how to act on it. The purpose of the data strategy in the company like Unity Technologies should be to serve as the best practice to manage data.

Even though the data strategy framework has been developed for the single case study, it could still be applicable the for other companies that are similar to the case study, i.e., fast-growing companies that have reached the momentum and are still growing. Fast-growing companies should have similar data issues in the company, as they started as startups, with their own in-house build systems that are very hard to scale when the company starts growing. Also, the focus of fast-growing companies is not on the functional strategies or procedures, but on the revenue increase. So the data strategy applicability might suit other fast-growing companies as the strategy is flexible and adaptable with all the needs of all the departments.

8.2 Limitations

In this part the limitations of the research will be discussed and how it might have impacted the outcomes of the research.

To begin with, this Master Thesis uses a single case study research approach. By investigating several companies within the industry and having a bigger picture, the data strategy framework could have been more applicable to the other companies. By investigating only one company it gives the depth of the problem, but the solution might not be widely applicable to the other companies.

The second limitation is that the respondents of the interviews, the managers of different departments within the company had their own agenda and bias towards the data. That might have impacted the findings and outcome of the data strategy. The main contact and stakeholder from the company, who has also validated the data strategy framework, could have been biased toward his own situation in the company and relationship with other departments. If there would have been a group discussion and a group interview, the findings, and the outcome might have been more objective.

8.3 Conclusion

This research describes the way of how to build a comprehensive data strategy in the organization. It discusses the problem area of data and strategies, how the data became so important that it needs a separate strategy in the organization. To support the idea of a comprehensive data strategy, different strategy alignment theories have been investigated which later became a foundation of the data strategy framework. The research has also investigated five different data strategy frameworks, which later were combined into the final conceptual data strategy framework. The big part of this research was empirical data, several interviews have been conducted with the senior managers in the company of the case study. The interviews helped to customized the data strategy framework and apply it to the case study.

What components should data strategy include?

In the introduction chapter, the concept of strategy has been discussed and what should be included in the strategy. In summary, a strategy is a plan or a framework, that includes long-term goals and objectives to strengthen the company in the market by considering its capabilities. The strategy includes and starts with vision, mission, and objectives then goes to the strategy methods of how to reach the objectives and finally comes to the detail action plan that is needed in order to reach those objectives. The data strategy should include objective, scope, and advantage that make up the data strategy statement.
Having the elements of the data strategy, the alignment with other strategies in the company can be performed in regards to those elements.

What are the main data challenges in the company?

This question has been raised to find out the main data issues that companies are having so that it could be solved with the data strategy. The data challenges that the companies have in general are described in the introduction chapter and the specific data challenges for the case study have been analyzed in the case study chapter. Some of the challenges overlap with the general ones, but other challenges are specific for the case study and possibly for the fast-growing companies. The causes of the data issues are even more specific for the case study, such as the in-house build systems and incoherent architecture, immature data processes and the culture does not focus on processes and procedures. These reasons and challenges make up the specific case of fast-growing companies.

To answer the subquestion, the main challenges in the company are: 1) poor quality of the source data - this is the starting point of the quality of data being transported to other systems and the quality of reporting; 2) there is no coherent data strategy in the organization - due to the lack of data leadership across the company Unity Technologies does not have overall data strategy, which leads to different data decisions between departments and accumulating data issues; 3) data architecture - there is no coherent data architecture across the company, different teams are building their own platforms which then causes issues with data integration.

How to align data strategy with business strategy and other strategies in the company?

The Theories chapter discusses strategy alignment theories and models, which helped to find the concept of how to align data and business strategy objectives. Three different theories have been investigated and it was decided to use the conceptual framework because some theories provided a different perspective on the alignment and only the theories were not enough for the data strategy framework.

The Three Levels of Strategy suggests that any strategy in the organization has to be aligned hierarchically, with the corporate mission and business objectives. That way the data strategy would be going in the same direction as the business. SAM model suggest the alignment between the functional strategy domains and the domain within the strategy itself. When using this concept, data strategy would be a comprehensive solution across all the organizations. The last Co-evolutionary IS/Business Alignment Model adds to the conceptual framework the part of alignment between the individual user and the system or in the case of data strategy, between the end user and data. Using this concept, it is ensured that the data addresses the needs of all the users and customers.

Combining all the three theories into the conceptual framework, data strategy alignment covers all the level in the organization, both hierarchically and horizontally/functionally, creating a comprehensive solution for data strategy.

What are the existing frameworks of data strategy? How to apply them to the case study?

In the literature review chapter, five different data strategy frameworks have been investigated. Most of the frameworks have the specific steps suggested that helps to develop a data strategy. Those steps consider what action needs to be taken to create a strategy. The other frameworks discuss the approach, the mindset that a company needs to figure out and follow in order to develop a data strategy. The frameworks have been described in detail and part of

them have been selected for the conceptual data strategy framework for the case study. The existing frameworks were not fully applicable to the case study and did not have enough significance to the case study and its challenges with data. Therefore, it was decided to use the conceptual framework, where investigated frameworks have been applied by the applicability to the case study. By applying the concepts from different data strategy frameworks, the most suitable for the case study are taken and build into the conceptual framework that could have the significance and real applicability to the case study company.

Problem formulation: how to build a comprehensive data strategy in the organization?

This main research question and the problem formulation is answered with the support from all the subquestions answered above. The answer to this question could also be illustrated by the data strategy graphical framework. For constructing this conceptual framework, strategy alignment theories and models have been investigated and concepts applied to the data strategy framework. The concepts that allowed data strategy to become a comprehensive solution across the company. From creating the awareness of data and gaining support from the top management to increase the data value in the company, a seven-step process has been developed. The applied concepts of existing data strategy framework, help to develop a process that the company has to go through in order to have a comprehensive data strategy solution.

References

[1] “Organizational challenges Disruptive business models Enhanced decision making THE AGE OF ANALYTICS: COMPETING IN A DATA-DRIVEN WORLD,” 2016.

[2] Stan Christiaens, “Four Strategies to Drive Real Business Value through Data Governance | Transforming Data with Intelligence,” 2016. [Online]. Available: https://tdwi.org/articles/2016/06/27/business-value-data-governance.aspx. [Accessed: 02-Sep-2018].

[3] “Gartner Survey Shows Organizations Are Slow to Advance in Data and Analytics,” 2018.

[4] L. Dallemule and T. H. Davenport, “ARTICLE ANALYTICS What’s Your Data Strategy?”

[5] M. Fleckenstein and L. Fellows, *Modern Data Strategy*. Cham: Springer International Publishing, 2018.

[6] “Setting the Data Strategy for Your Growing Organization.” [Online]. Available: https://www.stitchdata.com/resources/setting-your-data-strategy/. [Accessed: 02-Aug-2018].

[7] M. Dickey, “The G.O.A.T. may be a zebra | TechCrunch,” 2018. [Online]. Available: https://techcrunch.com/2018/04/15/the-g-o-a-t-may-be-a-zebra/. [Accessed: 02-Sep-2018].

[8] “Big data Changing the way businesses compete and operate Insights on governance, risk and compliance,” 2014.

[9] M. van Rijmenam, “What Data Do The Five Largest Tech Companies Collect - Infographic,” 2017. [Online]. Available: https://datafloq.com/read/what-data-do-the-five-largest-tech-companies-colle/427. [Accessed: 02-Sep-2018].

[10] W. W. ECKERSON, “How to Create a Data Strategy - Part I,” 2017. [Online]. Available: https://www.eckerson.com/articles/how-to-create-a-data-strategy-part-1-overview. [Accessed: 02-Sep-2018].

[11] J. W. Ross, I. Sebastian, and N. ; Fonstad, “Define Your Digital Strategy—Now « Center for Information Systems Research - MIT Sloan School of Management,” *MITCISR*, 2015.

[12] H. Mintzberg, “Patterns in Strategy Formation,” *Manage. Sci.*, vol. 24, no. 9, pp. 934–948, May 1978.

[13] N. Ritson, *Strategic Management*, no. 1. 2011.

[14] *Robson, 1997; Ward and Peppard, 2002*. .

[15] G. S. Nelson, “Developing Your Data Strategy: A practical guide.”

[16] S. Mankowitz and A. D. Snell, “Strategic and Financial Planning for Clinical Information Systems,” in *Clinical Informatics Study Guide*, Cham: Springer International Publishing, 2016, pp. 415–434.

[17] “Developing a Strategic Plan | Section 1. An Overview of Strategic Planning.” [Online]. Available: https://ctb.ku.edu/en/table-of-contents/structure/strategic-planning/vmosa/main. [Accessed: 02-Sep-2018].

[18] S. Kempe, “DataEd Slides: Implementing Successful Data Strategies - Developing Organizational Readiness and Framework - DATAVERSITY,” 2016. [Online]. Available: http://www.dataversity.net/dataed-slides-implementing-successful-data-strategies-developing-organizational-readiness-framework/. [Accessed: 02-Sep-2018].

[19] “Leganza, G., “Changing Your Approach to Information Strategy? You’re Not Alone,” Forrester Research, Inc., October 2, 2014.”

[20] “The 2016 global data management benchmark report.” [Online]. Available: https://www.edq.com/globalassets/white-papers/2016-global-data-management-benchmark-report.pdf. [Accessed: 02-Sep-2018].

[21] “Becoming an analytics-driven organization to create value A report in collaboration with Nimbus Ninety,” 2015.

[22] “The state of data quality - An Experian Data,” 2013.

[23] D. Loshin, “ADDRESSING FIVE EMERGING CHALLENGES OF BIG DATA.”

[24] S. Chen, “What Are The Stages of a Business Lifecycle and Its Challenges?” [Online]. Available: https://smallbusiness.yahoo.com/advisor/stages-business-lifecycle-challenges-100036770.html. [Accessed: 02-Sep-2018].

[25] N. Petch, “The Five Stages Of Your Business Lifecycle: Which Phase Are You In?,” 2016. [Online]. Available: https://www.entrepreneur.com/article/271290. [Accessed: 02-Sep-2018].

[26] C. Gnanasambandam, A. Miller, and K. Sprague, “Grow fast or die slow: The role of profitability in sustainable growth | McKinsey,” 2017. [Online]. Available: https://www.mckinsey.com/industries/high-tech/our-insights/grow-fast-or-die-slow-the-role-of-profitability-in-sustainable-growth. [Accessed: 02-Sep-2018].

[27] L. Matney, “With new realities to build, Unity positioned to become tech giant | TechCrunch,” 2017. [Online]. Available: https://techcrunch.com/2017/05/25/with-new-realities-to-build-unity-positioned-to-become-tech-giant/?guccounter=1. [Accessed: 02-Sep-2018].

[28] G. Da Silveira, D. Borenstein, and F. S. Fogliatto, “Mass customization: Literature review and research directions,” *Int. J. Prod. Econ.*, vol. 72, no. 1, pp. 1–13, 2001.

[29] R. V. Labaree, “Research Guides: Organizing Your Social Sciences Research Paper: Types of Research Designs.”

[30] M. Saunders, P. Lewis, A. Thornhill, S. • Lewis, and • Thornhill, “Research methods for business students fi fth edition.”

[31] R. K. Yin, *Case study research : design and methods*. .

[32] B. Alraggas and A. Alzayed, “The Alignment of Information Technology and Business Strategy in the Kuwaiti Companies,” 2014.

[33] N. Golafshani, “Understanding Reliability and Validity in Qualitative Research,” *Qual. Rep. 84*, vol. 8, no. 4, pp. 597–607, 2003.

[34] L. A. Guion, D. C. Diehl, and D. Mcdonald, “Triangulation: Establishing the Validity of Qualitative Studies 1.”

[35] V. A. Thurmond, “The point of triangulation,” *J. Nurs. Scholarsh.*, vol. 33, no. 3, pp. 253–258, 2001.

[36] T. Zenger, “Start with a Theory, Not a Strategy,” 2014. [Online]. Available: https://hbr.org/2014/06/start-with-a-theory-not-a-strategy. [Accessed: 02-Sep-2018].

[37] A. Bharadwaj, O. A. El Sawy, P. A. Pavlou, and N. V. Venkatraman, “Digital Business Strategy: Toward a Next Generation of Insights.” 01-Jun-2013.

[38] M. . Smith and P. Liehr, “Attentively Embracing Story: a middle-range theory with practice and research implications,” *Res. Theory Nurs. Pract.*, 1999.

[39] E. J. Watch, “E conomics in P ractice : F ollow -U p ‘ Theory ’ and ‘ Models ’:

Terminology Through the Looking Glass," vol. 5, no. 1, pp. 91–108, 2008.

[40] R. N. Giere, "How Models Are Used to Represent Reality," *Philos. Sci.*, vol. 71, no. 5, pp. 742–752, Dec. 2004.

[41] S. Imenda, "Is There a Conceptual Difference between Theoretical and Conceptual Frameworks?," *J. Soc. Sci.*, vol. 38, no. 2, pp. 185–195, 2014.

[42] R. M. Grant, *Contemporary strategy analysis : text and cases*. .

[43] G. Johnson, R. Whittington, K. Scholes, D. Angwin, P. Regnér, and S. Pyle, *Exploring strategy : text & cases*. .

[44] J. Baker and D. Jones, "A Theoretical Framework for Sustained Strategic Alignment and an Agenda for Research," vol. 8, no. 16, 2008.

[45] A. Renaud, I. Walsh, and M. Kalika, "Is SAM still alive? A bibliometric and interpretive mapping of the strategic alignment research field," *J. Strateg. Inf. Syst.*, vol. 25, no. 2, pp. 75–103, Jul. 2016.

[46] R. Online *et al.*, "Aligning Business Strategy with IT Strategy from Business Model to Enterprise in Saudi Arabia Public Sector," 2017.

[47] W. Afandi and S. Arabia, "The Impact of Strategic IT-Business Alignment: Evidence from Saudi Private Small and Midsize Enterprises," 2017.

[48] J. C. Henderson and N. Venkatraman, "Strategic alignment: Leveraging information technology for transforming organizations," 1993.

[49] "Mithas & Lucas 2010; Mithas et al. 2013b; Oestreicher-singer & Zalmanson 2013; Grover & Kohli 2013b; Nylen & Holmstrom 2015; Teubner 2013; Benbya & McKelvey 2006)."

[50] H. Benbya and B. McKelvey, "Using coevolutionary and complexity theories to improve IS alignment: a multi-level approach," *J. Inf. Technol.*, vol. 21, pp. 284–298, 2006.

[51] "(Kearns & Lederer 2000; Tallon et al. 2000)."

[52] "Conceptual Framework What Do You Think Is Going On?" .

[53] "A guide for writing scholarly articles or reviews for the Educational Research Review." .

[54] M. J. Mazzei and D. Noble, "Big data dreams: A framework for corporate strategy," *Bus. Horiz.*, vol. 60, no. 3, pp. 405–414, May 2017.

[55] T. H. . Davenport, B. Paul, and Randy Bean;, "How 'Big Data' Is Different," 2012.

[56] M. Peterman, "The Truth about Data Truth: SSoT vs MVoT - CSG," 2017. [Online]. Available: https://www.csgsolutions.com/blog/truth-data-truth-ssot-vs-mvot/. [Accessed: 02-Sep-2018].

[57] "About Eckerson Group." [Online]. Available: https://www.eckerson.com/about. [Accessed: 02-Sep-2018].

[58] W. W. ECKERSON, "How to Create a Data Strategy: Part II – Ten Steps to Success," 2017. [Online]. Available: https://www.eckerson.com/articles/how-to-create-a-data-strategy-part-ii-ten-steps-to-success. [Accessed: 02-Sep-2018].

[59] W. W. ECKERSON, "How to Create a Data Strategy Part III: From Design to Execution," 2017. [Online]. Available: https://www.eckerson.com/articles/how-to-create-a-data-strategy-part-iii-from-design-to-execution. [Accessed: 02-Sep-2018].

[60] "Gartner Predicts 15 Percent of Successful Chief Data Officers Will Move to CEO, COO, CMO or Other C-Level Positions by 2020," 2016.

[61] M. Lakoju and A. Serrano, “A strategic approach for visualizing the value of big data (SAVV-BIGD) framework,” in *2016 IEEE International Conference on Big Data (Big Data)*, 2016, pp. 1334–1339.

[62] A. Laplante and B. Sharma, “Architecting Data Lakes.”

[63] J. O’brien and Ceo, “THE DEFINITIVE GUIDE TO THE.”

[64] N. Miloslavskaya and A. Tolstoy, “Big Data, Fast Data and Data Lake Concepts,” *Procedia Comput. Sci.*, vol. 88, pp. 300–305, Jan. 2016.

[65] M. Chessell, F. Scheepers, M. Strelchuk, R. Van Der Starre, S. Dobrin, and D. Hernandez, “Redguide The Journey Continues From Data Lake to Data-Driven Organization.”

[66] “Data Lake Experience.” [Online]. Available: http://www.elevondata.com/ElevonDataweb/services/data-lake. [Accessed: 02-Sep-2018].

[67] “The business value of a business glossary.”

[68] “Veritas Global Databerg Report Finds 85% of Stored Data Is Either Dark, or Redundant, Obsolete, or Trivial (ROT).” [Online]. Available: https://www.veritas.com/news-releases/2016-03-15-veritas-global-databerg-report-finds-85-percent-of-stored-data. [Accessed: 02-Sep-2018].

[69] “DATA STRATEGY DEVELOPING YOUR ROADMAP FOR DRIVING BUSINESS VALUE WITH DATA.”

[70] N. Zhang, “An Overview of Data Governance An Overview of Data Governance,” no. September 2016, 2017.

[71] T. C. Redman, *Data Driven: Profiting from Your Most Important Business Asset*. Harvard Business Review, 2008.

[72] “Leganza, G., ‘Drive Information Strategy Performance Management with Capability Models,’Forrester Research, Inc., October 2, 2014.”

[73] “Data protection in the EU | European Commission.” [Online]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en. [Accessed: 02-Sep-2018].

[74] *I Legislative acts REGULATIONS*. .

[75] “Data Protection Principles – Unlocking the EU General Data Protection Regulation | White & Case LLP International Law Firm, Global Law Practice.” [Online]. Available: https://www.whitecase.com/publications/article/chapter-6-data-protection-principles-unlocking-eu-general-data-protection. [Accessed: 02-Sep-2018].

[76] “Guidelines on the protection of personal data in IT governance and IT management of EU institutions,” 2018.

[77] “Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1,” Gaithersburg, MD, Apr. 2018.

[78] “Guide for applying the risk management framework to federal information systems : a security life cycle approach,” Gaithersburg, MD, Jun. 2014.

[79] “Intangible Property Valuation Guidelines | Internal Revenue Service.” [Online]. Available: https://www.irs.gov/irm/part4/irm_04-048-005. [Accessed: 02-Sep-2018].

[80] I. Builders, “How to Make Money or Save Money With Data and Analytics.”

[81] B. Bilbao-Osorio, S. Dutta, and B. Lanvin, “The Global Information Technology Report 2014 Rewards and Risks of Big Data,” 2014.

[82] S. Ravindra, “Data Monetization: A New Way of Thinking | Transforming Data with Intelligence,” 2017. [Online]. Available: https://tdwi.org/articles/2017/09/07/biz-all-data-monetization-new-way-of-thinking.aspx. [Accessed: 02-Sep-2018].

[83] K. International, “Framing a winning data monetization strategy,” 2015.

[84] “Hortonworks Big Data Maturity Model HORTONWO RKS BIG DATA MATURITY MODEL,” 2016.

[85] R. Cosic, G. Shanks, and S. Maynard, “Towards a Business Analytics Capability Maturity Model,” in *23rd Australasian Conference on Information Systems*, 2012.

[86] J. Becker, R. Knackstedt, and J. Poeppelbuss, “Developing Maturity Models for IT Management – A Procedure Model and its Application,” *Bus. Inf. Syst. Eng.*, vol. 1, no. 3, Jan. 2011.

[87] S. Myburgh, *Conceptualisation of a Big Data Maturity Model Based on Organisational ... - Suzanne Myburgh - Google Books*. 2015.

[88] “Info-Tech Research Group | Info-Tech Research Group.” [Online]. Available: https://www.infotech.com/research/ss/leverage-big-data-by-starting-small/it-big-data-maturity-assessment-tool. [Accessed: 02-Sep-2018].

[89] D. Maturity, “Data Maturity.”

[90] “Data & Analytics Maturity Assessment - Ironside - Business Analytics. Data Science. Information Management.” [Online]. Available: https://www.ironsidegroup.com/data-analytics-maturity/. [Accessed: 02-Sep-2018].

[91] “Assessing Community Needs and Resources.” [Online]. Available: https://ctb.ku.edu/en/table-of-contents/assessment/assessing-community-needs-and-resources/swot-analysis/main. [Accessed: 02-Sep-2018].

[92] “Unity Public Relations.” [Online]. Available: https://unity3d.com/public-relations. [Accessed: 02-Sep-2018].

[93] T. Wijman, “Global Games Market Revenues 2018,” 2018. [Online]. Available: https://newzoo.com/insights/articles/global-games-market-reaches-137-9-billion-in-2018-mobile-games-take-half/. [Accessed: 02-Sep-2018].

[94] D. Takahashi, “How Unity convinced investors it’s worth $2.6 billion | GamesBeat,” 2017. [Online]. Available: https://venturebeat.com/2017/05/25/why-unity-was-able-to-raise-400-million-at-a-2-6-billion-valuation/. [Accessed: 02-Sep-2018].

[95] “Unity Unveils 2018 Roadmap at GDC – Unity Blog.” [Online]. Available: https://blogs.unity3d.com/2018/03/20/unity-unveils-2018-roadmap-at-gdc/#unity2018. [Accessed: 02-Sep-2018].

[96] J. S. Saltz and I. Shamshurin, “Exploring the process of doing data science via an ethnographic study of a media advertising company,” in *2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 2098–2105.

[97] “EU GDPR Information Portal.” [Online]. Available: https://www.eugdpr.org/. [Accessed: 02-Sep-2018].

[98] “Scrum Guide.” [Online]. Available: https://www.scrumguides.org/scrum-guide.html. [Accessed: 02-Sep-2018].

[99] “Agile Teams – Scaled Agile Framework.” [Online]. Available: https://www.scaledagileframework.com/agile-teams. [Accessed: 02-Sep-2018].

[100] C. FORSBERG, “Beautiful Delivery Process - SogetiLabs,” 2018. [Online]. Available: http://labs.sogeti.com/beautiful-delivery-process/. [Accessed: 02-Sep-2018].

[101] “William Thomson, Lord Kelvin.”