Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Systems Engineering

Master's Thesis

Business Model for Data Governance Platform

Tia Semerádová

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

Bc. Tia Semeradova

Informatics

Thesis title

Business model for data governance platform

Objectives of thesis

This thesis aims to build a business model for a new data governance platform that connects metadata and extends and supports business intelligence across a company.

A. The First main objective is to create a business model for the new platform.

1a) First subpart of this objective contains:
1aa. a description of the solution proposed in the form of the platform analysis
1ab. an examination of existing solutions in the EU market and in the Word market.
2a) Second subpart aims to specify a project schedule for the platform development.

B. The second main objective is to define buyers' personas.

1b) A subpart of this objective is to conduct and analyse interviews with potential customers and according to Design thinking methodology.
2b) The second subpart of this objective will be to create use cases based on the user requirements.
3b) The third subpart of this objective will be to build a business case for the platform.

C. The third main objective will be to design a wireframe for the platform.

The Wireframe will be a part of the platform that distinguishes it from its competition, which is identified in the first objective. The Wireframe will also correspond with the users' requirements that will have been analysed in the second main objective.

Methodology

The theoretical part of this thesis will be based on a critical review of relevant articles and literature to analyse modern approaches to business models, design thinking, user case and the described platform environment.

The practical part will include business model creation, interviews and interview analyses using the design thinking method, cost analysis, and platform Wireframe.

The proposed extent of the thesis

60-80 pages

Keywords

business model, business case, design thinking

Recommended information sources

Hague, P. (2019). The Business Models Handbook (1st ed.). Kogan Page. (Original work published 2019)
Hamm, M. J. (2014). Wireframing Essentials: An introduction to user experience design: Learn the fundamentals of designing the user experience for applications and websites. Packt Publishing.
Oberholzer-Gee, F. (2021). Better, Simpler Strategy. Harvard Business Review Press. (Original work published 2021)
Osterwalder, A., & Pigneur, Y. (2013). Business Model Generation (1st ed.). Wiley. (Original work published 2013)
Revella, A. (2015). Buyer personas: How to gain insight into your customer's expectations, align your marketing strategies, and Win More Business. John Wiley & Sons.
Rognerud, J. (2014). Ultimate Guide to Optimizing Your Website . Entrepreneur Press. (Original work published 2014)
van der Pijl, P. van der, Lokitz, J., Wijnen, R., & Lieshout, M. van. (2020). Business Model Shifts (1st ed.). Wiley. (Original work published 2020)
Yayici, E. (2016). Design thinking methodology book. ArtBizTech.

Expected date of thesis defence

2022/23 SS – FEM

The Diploma Thesis Supervisor

Ing. Petra Pavlíčková, Ph.D.

Supervising department

Department of Systems Engineering

Electronic approval: 26. 11. 2022

doc. Ing. Tomáš Šubrt, Ph.D.

Head of department

Electronic approval: 28. 11. 2022

doc. Ing. Tomáš Šubrt, Ph.D.

Dean

Prague on 03. 03. 2023

Declaration

I declare that I have worked on my master's thesis titled " Business model for data governance platform" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the master's thesis, I declare that the thesis does not break any copyrights.

In Prague on 22.03.2023 _____

Acknowledgement

I would like to thank Ing. Petra Pavlíčková, Ph.D. for all her advice and support during my work on this thesis. I would also like to thank all the respondents who had participated in this thesis's qualitative research.

Business Model for Data Governance Platform

Abstract

The primary goal of this master's thesis is to develop a business model for a new data governance platform, T.D.G.P. The author has defined several key objectives will be accomplished in later chapters, such as identifying and analysing existing solutions in the market. The second aim is to define buyer personas by conducting interviews with potential customers and creating use cases based on their needs. The third objective entails crafting a wireframe platform that meets user requirements while distinguishing itself from competitors.

The theoretical part of this study is based on critical exploration of relevant literature, covering modern approaches to business models, design thinking, user cases, and platform ecosystems. This literature review sets the foundation for constructing a business model incorporating the field's most recent trends and best practices.

The practical aspect of this research involves performing market and strategic analyses of the current market and interviewing potential customers to discern their preferences and requirements. Using the design thinking approach, these interviews will be analyzed to uncover potential users' underlying needs and motivations. Based on the insights from these interviews, use cases and a business case will be developed to support the platform's conceptualization. This thesis will also propose a wireframe model designed to cater to user requirements and differentiate the platform from competitors. The wireframe will visually showcase the platform's features and functions.

Keywords: Data Governance Platform, Business Model, Metadata Management, Business Intelligence, Design Thinking, Buyer Personas, Use Cases, Market Analysis

Business model pro Platformu pro správu dat

Abstrakt

Hlavním cílem této magisterské práce je vytvořit obchodní model nové platformy pro správu dat T.D.G.P. Autor si stanovil několik klíčových cílů, které budou splněny v dalších kapitolách, jako je identifikace a analýza stávajících řešení na trhu. Druhým cílem je definovat persony kupujících provedením rozhovorů s potenciálními zákazníky a vytvořením případů užití na základě jejich potřeb. Třetí cíl zahrnuje vytvoření drátěného modelu platformy, který splňuje požadavky uživatelů a zároveň se odlišuje od konkurence.

Teoretická část této studie je založena na kritickém zkoumání relevantní literatury, která zahrnuje moderní přístupy k obchodním modelům, designovému myšlení, případům uživatelů a ekosystémům platforem. Tento přehled literatury vytváří základ pro konstrukci obchodního modelu zahrnujícího nejnovější trendy a osvědčené postupy v této oblasti.

Praktický aspekt tohoto výzkumu zahrnuje provedení tržní a strategické analýzy současného trhu a rozhovory s potenciálními zákazníky s cílem zjistit jejich preference a požadavky. S využitím přístupu designového myšlení budou tyto rozhovory analyzovány s cílem odhalit základní potřeby a motivace potenciálních uživatelů. Na základě poznatků z těchto rozhovorů budou vypracovány případy užití a obchodní případ na podporu konceptualizace platformy.

V této práci bude také navržen model wireframe, který má vyhovět požadavkům uživatelů a odlišit platformu od konkurence. Drátěný model bude vizuálně prezentovat vlastnosti a funkce platformy.

Klíčová slova: Platforma pro správu dat, obchodní model, správa metadat, business intelligence, designové myšlení, persony kupujících, případy užití, analýza trhu.

Table of content

1 Introduction

In the current data-centric business landscape, the significance of data management is indisputable. Organizations must administer and oversee their data assets while ensuring data quality and compliance and minimizing the risk of data breaches or leaks. Effective data management allows organizations to derive maximum value from their data and make well-informed business decisions. However, the market's existing state of data management could be enhanced as numerous organizations grapple with managing their data efficiently. Among available solutions that aim to solve the data problem are data governance platforms.

Developing a successful data governance platform requires a well-defined business model. A comprehensive business model must consider the organization's distinct needs and objectives, as well as the platform's capabilities and constraints. The model should explicitly outline the platform's value proposition, target audience, revenue sources, and cost structure.

To devise such a business model, an exhaustive market analysis is essential. This encompasses determining the platform's target audience, evaluating the extent of competition, and performing a SWOT analysis to identify the platform's strengths, weaknesses, opportunities, and threats. A PEST analysis is also needed to examine the political, economic, social, and technological factors that may impact the platform's success in the market. Porter's five forces model can also be employed to gauge the market's competitive strength and ascertain the platform's competitive edge.

Utilizing design thinking interviews to gather data and insights during business model development, the researcher can better understand the company's target market's needs and preferences. Qualitative interviews can offer valuable perspectives on how users interact with data and make decisions based on that data.

A data governance platform offers a cohesive approach to data management and allows organizations to establish policies, processes, and standards for data usage and quality. This is particularly crucial as data has become a liability that must be suitably managed and regulated; otherwise, it can result in inefficiencies, errors, and risks. It is also an underappreciated asset that, when leveraged properly, can considerably benefit the company in various aspects.

The practical portion of this thesis is devoted to formulating a business model for a data management platform. This entails conducting an in-depth market analysis, crafting a unique value proposition for the platform, and designing a project plan and timeline. Moreover, buyer personas are established based on and then employed through interviews with potential customers. Use cases are demonstrated to understand the platform's functionality and user experience clearly. Finally, a wireframe is developed that visualizes the platform's unique value proposition.

1.1 Objectives

The primary goal of this thesis is to develop a comprehensive business model for a new data governance platform which connects metadata and enhances business intelligence throughout an enterprise. This primary objective is further divided into two distinct sub-objectives.

The first sub-objective is to create a detailed business model for the proposed platform. This involves analysing the solution, examining existing data governance platforms in the global market, and specifying a project schedule for the platform's development.

The second sub-objective is to define the target audience as buyers' personas for the data governance platform. This involves conducting and analysing interviews with potential customers, following the principles of design thinking methodology, creating use cases based on user requirements, and developing a business case for the platform.

The third primary objective is to design a wireframe for the platform. The wireframe will be an integral part of the platform that distinguishes it from its competitors and meets users' requirements, as identified in the second main objective.

1.2 Methodology

The theoretical component of this thesis will entail a comprehensive analysis of related articles and literature aimed at scrutinizing contemporary methodologies and paradigms regarding business models, design thinking, user case, and platform environments. A critical review will be conducted to get a deep understanding of these domains and how they influence the development of a data governance platform.

The practical part of this thesis will aim at creating a business model for the data governance platform. To develop a comprehensive business model, marketing and strategic analysis will be conducted, together with a proposed business case that would provide more theoretical to practical implementation. This will, also contain cost calculations for the project based on the information from the portal ISPV from the Czech Ministry of Labour and Social Affairs. The wage costs will be calculated based on this information to ensure a realistic financial plan for the platform's

implementation. Furthermore, interviews will be conducted with anonymized respondents who work at large corporations with at least 200 employees. These interviews will later be analysed using design thinking methods to gain insights into the users' needs and preferences. Based on those interviews, a wireframe of a distinguishing feature of the data governance platform will be proposed.

Literature Review

1.3 Introduction

The literature review of this thesis explores the current state of knowledge in business models, design thinking, marketing analysis and data governance. These topics are followed up on in the practical part of this thesis. The practical part then which aims to propose a business model for a data governance platform.

The first section of the literature review, Business Model Historical Development, explores the evolution of the business model concept throughout history. The author of this thesis examines the frequency and context in which the term has been used, as well as the evolution of its definition over time. This section of this thesis also provides an overview of the various business models proposed in the literature, focusing on the following models: Business Model Canvas (BMC), Value Proposition Design, Strategy Sketch, and Lean Canvas.

The second section of the literature review, Design Thinking, focuses on utilizing design thinking in business model development. The author examines the use of design thinking interviews to gather data and insights while developing a business model.The Market Analysis section in this thesis stresses the importance of conducting market research to develop a successful business model. SWOT analysis, PEST analysis, and Porter's Five Forces Model are used to describe the external environment, identify opportunities and threats, and inform decision-making. The section also explores the TOWS matrix and A'WOT method, which build upon SWOT analysis. Porter's Five Forces Model evaluates the collective strength of five forces that impact industry profitability and is essential for strategic planning and decision-making.

The last section of the literature review explores the various aspects of developing a business model for a data governance platform. The author describes data quality, metadata, data access, and data life cycle management and how they can be integrated into the proposed business model. The section also provides an overview of current best practices and challenges in the field of data governance and how they can inform the development of a business model for data governance.

1.4 Business model historical development in the scientific literature

The term business model is a relatively new phenomenon. Its most significant increase in popularity occurred around the early 21st century. According to Carlos M. DaSilva and Peter Trkman (DaSilva and Trkman 2014), "*the widespread use of the business model terminology seems intrinsically connected with technology-based companies*." Christian Nielsen and Morten Lund state that in the late 1990s, the term business model had almost the same meaning as e-business for how tightly connected those two terms were. (Nielsen and Lund 2018)

1.4.1 Business model – the origin

No single source or article would be wildly considered to be the origin of the business model concept. The first mention of a "business model" was in an academic paper written by Bellman, Clark, Malcolm, Craft, and Ricciardi in 1957. (Bellman et al. 1957) The term in this article is mentioned only once on page 474 in the context of creating mathematical problems and simulations to solve business problems. Since the term is not used with the same meaning that it is understood nowadays, the article is generally not considered to be the source of the first idea of a business model.

Multiple authors ((Wirtz et al. 2016) (DaSilva and Trkman 2014) (Rocha et al. 2018)) agree that the initial use of the term in 1960 in an academic article by Jones (Jones 1960), however the term itself was still not used in the definition that is understood today as the full name of the article reads: *Educators, electrons and business models: problem in synthesis* and the article itself focuses more on education in business rather than business structure and development. Even though the term business model is used in the article title, it is not mentioned in the text.

1.4.2 The trend of scientific articles discussing the business model

Osterwalder researched the number of articles that used business models in their title, abstract, keywords and full text and found that the term experienced a boom in popularity after the 1990s. (Osterwalder et al. 2005) Carlos M. DaSilva and Peter Trkman connect this phenomenon to developing information and communication technologies. (DaSilva and

Trkman 2014) They also researched this topic with an expanded focus on the first mention of the business model in an academic article in 1957. For their research, they used the Web of science. As an explanation for their choice of platform, they refer to the report by Norris and Oppenhim, where they say that the Web of science "offers a reliable coverage and historical overview at the journal, article and cited-reference level". Their research found relatively few articles using the term business model in any capacity until 1992. The term then was slowly gaining popularity until it encountered rapid growth in 1998, and this trend continued through 2010, which was used as an upper limit in the article research.
The author of this diploma thesis conducted the same research via the portal Scopus. Scopus was used because it is the most accessible portal with scientific articles available to university students in the Czech Republic.

The research examined articles with business models in their Key Words, Abstract and Title. However, since the earliest article available on this platform is from 1968, it is essential to point out that the graph below does not show an accurate number of articles but only serves as a form of general visualization of grow of articles that use the term till the year 2021.
The years 1968 to 1990 are not included since there was a very low number of articles.

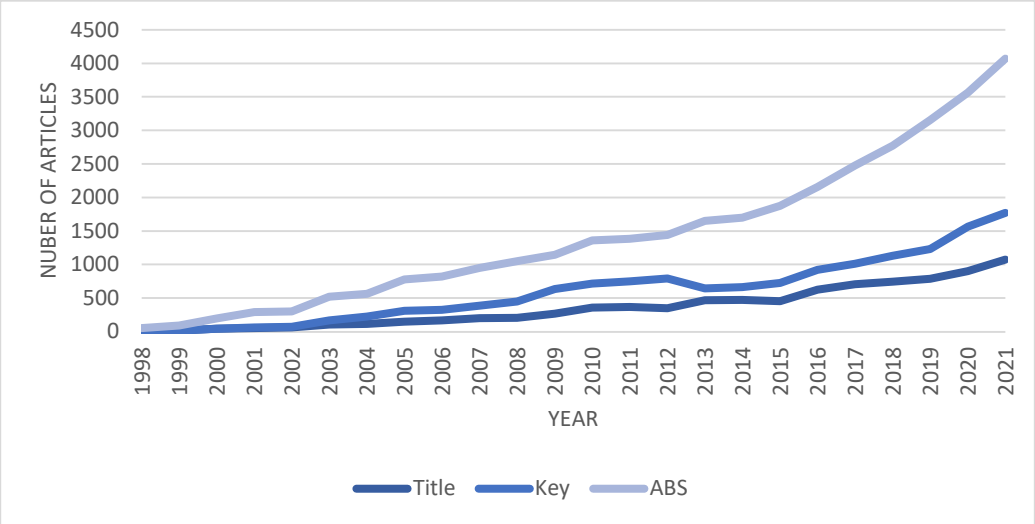


Figure 1 Source: Elsevier Scopus database

As we can see on the chart, the trend continues to rise exponentially even after 2010. The most significant increase in escalation can be seen in articles containing the term business model in their abstract.

1.4.3 The historical meaning of business model in the scientific literature

According to Christian Nielsen and Morten Lund, the business model concept started receiving a much more general meaning in management literature than the e-biz rhetoric around 2001-2002. (Nielsen and Lund 2018) They cite an article from Magretta that appeared in Harvard Business Review in which the author states that creating a business model is a lot like writing a new story. She also compares the business model to the scientific method in an organizational context. She describes an aspect of value to the business model and states that the business model also comes with the question of who the customer is and what they value. (Joan Magretta 2002) In other words, a business model's original purpose is to communicate complex business ideas to potential investors. (Geissdoerfer et al. 2020)

After the tech stock crash, the business model's focus changed, at least from the point of investors. As before, the general idea was that everything containing the " dot com" in their mind frame would eventually turn out profitable.

Christian Nielsen and Morten Lund created a table in which they mapped the literature overview of the business model research field between the years 1997 to 2013. They pointed out three main fields in which scientific articles discussing business models are oriented. (Wirtz et al. 2016)

- Technology-oriented
- Organization theory oriented
- Strategy oriented

With this table, they described the shift in article topic focuses as technology-oriented before 2002, especially in electronic business. From 2002 they say that articles were more strategy-oriented even though technology-oriented articles remain. They show organization-oriented articles, such as from

authors Treacy Wiersema and Lindrer Cantrell. However, the authors point out that "*in comparison with the other two currents in scientific discourse, they play a subordinate role*". The clear segregation of the theme of articles seems to be retreating in recent years, and there are approximately a similar number of articles in each school of orientation. (Wirtz et al. 2016)
Nielsen and Lund then describe the differences between how each school of orientation understands the business model concept.

Technological orientation

They write that authors of technological orientation see business models much more abstractly and, in a sense, as a representation of a company, in contrast to very early days authors with technological orientation who saw the business model as a small part of the company. (Wirtz et al. 2016)

Organizational orientation

Authors of organization orientation shared the view with authors of technological orientation, in that the business model is a strongly abstract tool to provide a picture of a company's competitive situation. (Wirtz et al. 2016)

Strategy orientation

In the strategy-oriented view, the authors point out the increasing consensus among authors, especially strategy-oriented authors, who focus on the purpose of the business model concept and the role within already existing business concepts. From this school of view, a new question is formed: What is the difference between a business model and strategy? (Wirtz et al. 2016)
Casadesus-Masanell and Ricart answer this question by explaining that strategy and business model, though related, are different concepts: a business model is the direct result of strategy but is not itself a strategy. (Casadesus-Masanell and Ricart 2010)

Later, many authors pointed out that the term business model still lacks a clear definition. For example, Teece points out that the "*study of business models is an interdisciplinary topic which has been neglected*" then later in the same article, in which he aims to help to remedy the definition of business model, he states that *"business models are frequently mentioned but rarely analyzed:*

therefore, they are often poorly understood". In the article's conclusion, he describes how organizations consciously or unconsciously use a specific business model.

The Teeces business model describes the design or architecture of the value creation, delivery, and processes used in the business. The core of a business model is how it maps the needs and financial capacity of the customer, establishes how the company responds to and provides value to the customer, persuades the customer to pay for value, and turns that payment into profit through the effective setup and operation of the various value chain elements.

To summarize Teece's view, a business model reflects management's assumptions about what customers want, what they want, what they are willing to pay for it, and how an organization might suit those demands best. (Teece 2010)

1.4.4 The modern definition of business model in the scientific literature

Even after many authors have discussed the lack of clear meaning of business model and its distinguishing from other business terms, there still seems to be a need for a unifying definition used in the scientific literature; the public literature would also accept that.

Rocha, Perez, Romero and Numez propose a business model definition in their article. They add that together with the definition. They also propose a table of elements to include elements they believe would bring importance to the future of organizations. (Rocha et al., 2018)

In their research, they examined 45 studies. Their research focused on identifying the elements that constitute the business model. (Rocha et al., 2018) In their table, they summarized a list of 30 elements. They found that the most relevant elements in the studies were Value Proposition (with 69% of articles identifying them), Revenue Streams (with 51% of articles identifying them and Customers (40% of articles identifying them).

They introduce a new definition which states, "*A business model is a story about an organization's value proposition, defined through a panel of critical elements that comprise its way of operating, and based on three fundamental characteristics:*

• *It must be unique to guarantee the generation of competitive advantages and that customers will choose it in the short term*
• *It must be adaptable to guarantee its long-term viability*
• *It must be scalable to ensure that revenue grows above investment in resources*" (Rocha et al. 2018)

1.5 Business models

There are many different business models, each with unique advantages and disadvantages. Understanding and choosing a suitable business model can define and dictate a way for a company's success. This section explores some of the most popular business models used in today's business world and discusses the pros and cons of each.

The Business Model Canvas (BMC) is a widely used tool for visualizing and designing a company's or product's value proposition, infrastructure, consumers, and finances. Developed by Alexander Osterwalder in 2005, the BMC has become a popular tool for entrepreneurs and businesses looking to align their efforts and identify trade-offs. However, there are also criticisms of the BMC, with some arguing that it does not consider competition, business goals, or key performance indicators. In response to these criticisms, Osterwalder has developed the Value Proposition Design, which aims to make the value more understandable and optimize the product simultaneously. Additionally, Jeroen Kraaijenbrink has proposed the Strategy Sketch as an updated version of the BMC that incorporates strategy elements without increasing complexity. This thesis section will examine these models' benefits, drawbacks, and potential use cases.

1.5.1 Business model- Canvas (BMC)

The Business Model Canvas provides a visual chart detailing a company's or product's value proposition, infrastructure, consumers, and finances. Highlighting potential trade-offs helps firms align their efforts. (Taipale-Erävala, Salmela, and Lampela 2020)
Alexander Osterwalder first presented the nine "building blocks" of the business model design template that eventually became known as the Business Model Canvas in 2005, based on his earlier work on business model ontology. Since Osterwalder's pieces were first introduced around 2008, more canvases targeting particular markets have appeared. (Osterwalder and Pigneur 2010)

There are many benefits arising from using BMC. Hong Yuh Ching, in his article, includes a list of benefits, which talks about the advantages of a model based on practice and easily applicable to rising entrepreneurs. (Ching 2013) However, he follows with a list of potential disadvantages of implementing the BMC. Ching claims that the following five points arise with the side of the medal of simplicity.

The list of five negatives is written below:

1. *No broad analyses of competition*
2. *No, taking into account competition structures (and, therefore, potential synergy*
3. *effects)*
4. *No formulating of business goals*
5. *No, taking into account KPIs and performance measurements applicable for innovation, not so much for transforming existing models* (Ching 2013)

Kraaijenbrink introduces his list of advantages and disadvantages of business model canvas and proposes a canvas that fills in some gaps. (Jeroen Kraaijenbrink 2022)
He states that Business Model Canvas incorporates all of the essential components of a business model on a single piece of paper, makes it distinctive and serves as its primary advantage. However, he then moves on to talk about the drawbacks of the business model canvas and names key elements that he finds are missing.

Competition, Values, and Goals are, according to Kraaijenbrink, missing the Business Model Canvas and the Business Model Generation book in which it is introduced, even though those elements are present in strategy textbooks.

He also criticized some elements present in the business model canvas. "Customer relationships" and "Channels" elements were, according to him, left blank or only contained relatively trivial or operational content. Furthermore "Key Activities" element, in his opinion ", seemed more related to how to execute or implement a business model rather than that it was part of the business model itself." (Jeroen Kraaijenbrink 2022)

1.5.2 Value Proposition Design

The Value Proposition Design is another model designed by Oserwalder to fill the need for Value definition in the business model canvas. He references the business model canvas in the book. This new model helps to make the value more understandable and helps optimize the product at the same time.

There are two main blocks: customer profile and a company's value proposition. (Alex Osterwalder et al. 2014)

Customer Profile (Alex Osterwalder et al. 2014)

- Gains
- Pains
- Customer jobs

Value Map (Alex Osterwalder et al. 2014)

- Gain creators
- Pain relievers
- Products and services

The main objective of Value Proposition Design is to provide a tool to the reader that will help him identify whether there is a fit between the new product and the market, but also whether there is a need to refine the existing product to remain a good fit for the current market. (Alex Osterwalder et al. 2014)

1.5.3 The Strategy Sketch

The Strategy Sketch was proposed as an update for BMC by Jeroen Kraaijenbrink. The model aimed to offer a complete version of BMC regarding its coverage of Strategy elements without making the model more complex. (Jeroen Kraaijenbrink 2022)
The strategy sketch consists of ten elements described in more detail in Kraaijenbrink's book.

The elements are listed below.

1. ***Resources and competencies***. *What you have, what you are good at, what makes you unique.*
2. ***Partners***. *Whom do you work with, and who makes your products or services more valuable.*
3. ***Customers and needs.*** *The organizations and people you serve and which of their needs you fulfil.*
4. ***Competitors.*** *Other that your customers will compare you to in deciding whether or not to buy your products or services.*
5. ***Value proposition***. *What products and services, how you offer them, and what added value they have for the customer.*
6. ***Revenue model***. *What you receive in return for your offer, from whom, how and when.*
7. ***Risk and costs***. *What financial social and other risks and costs you bear and how you manage there.*
8. ***Values and goals***. *What you want, where you want to go, and what you find important.*
9. ***Organizational climate***. *What your culture and structure look like and what is special about them.*
10. ***Trends and uncertainties.*** *What happens around you affects your organization and what uncertainties you face.*

1.5.4 Lean Canvas

The Lean Canvas is a tool that helps entrepreneurs and business leaders design, plan and test their business models. The Lean Canvas is based on the Business Model Canvas. It is a visual representation of a business model that helps companies to understand and communicate their value proposition, customer segments, channels, revenue streams, and key resources and activities efficiently (Maurya Ash 2022)

Lean Canvas helps companies to validate their business model before committing significant resources to it. By testing the assumptions behind each business model element, companies can quickly identify which parts of the model are working and which are not and adjust accordingly (Maurya Ash 2022). This approach aligns with the Lean Startup methodology, which emphasizes the need to validate business models through customer feedback and experimentation. (Alex Osterwalder et al. 2014)

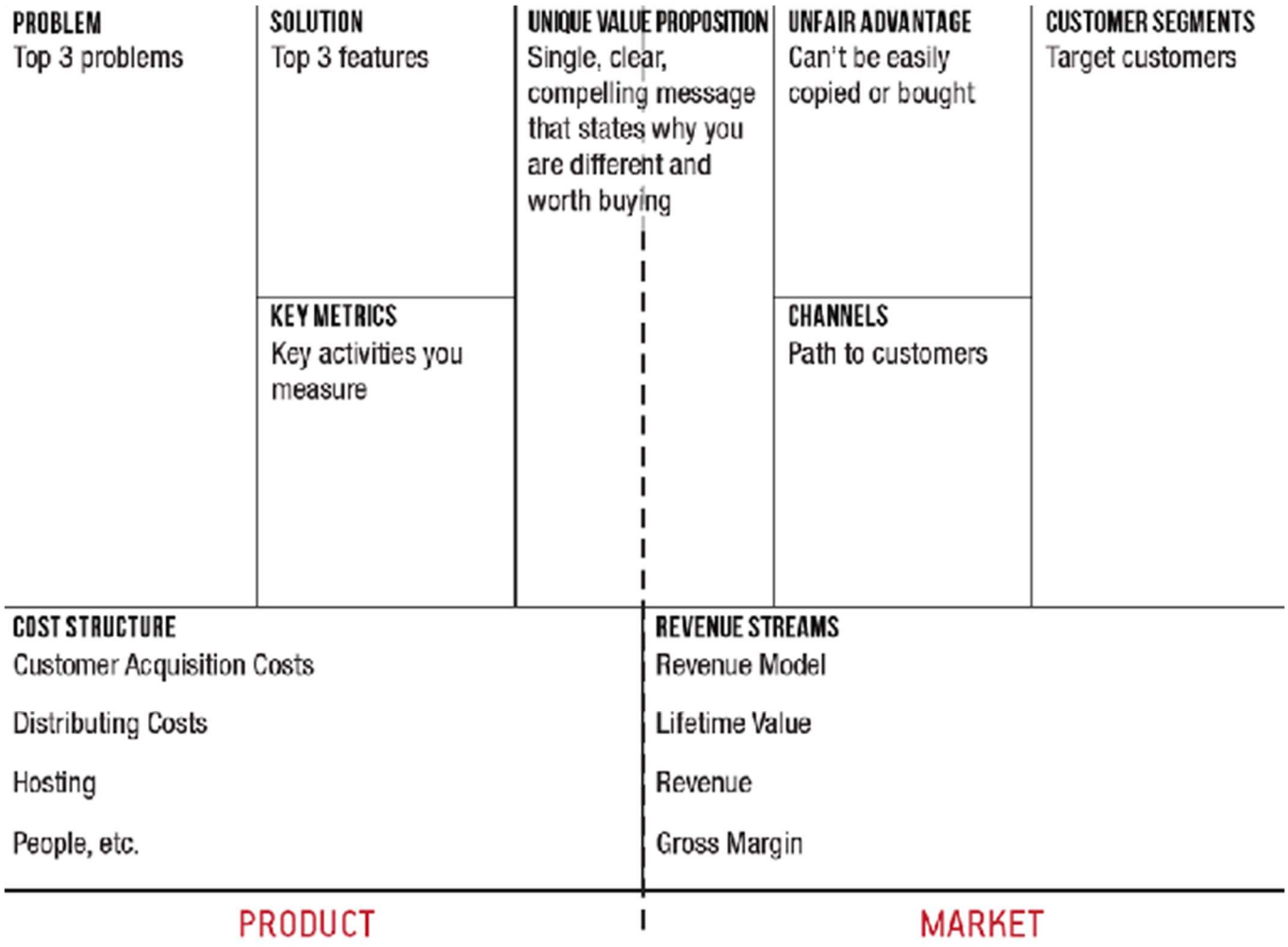


Figure 2 Lean Canvas (Maurya Ash 2022)

The Lean Canvas also helps startups to focus on the most critical aspects of their business and save time and resources on activities that do not contribute to their overall strategy. By identifying and prioritizing the key risks and uncertainties in their business model, startups can use Lean Canvas to develop a clear and actionable plan for testing and validating their assumptions (Maurya Ash 2022)

Lean Canvas is a flexible and adaptable tool that companies can use at different stages of development. Early-stage startups can use it to validate their initial business model, as well as established companies to develop new products or services or to explore new market opportunities (Maurya Ash 2022)

1.6 Market Analysis

When creating a business model, it is necessary to conduct a market analysis to understand the external environment in which the business will operate. Market analysis helps to identify the demand for a product or service, the target market and the competition. If the market analysis were excluded, the business model could lack context and lead to unrealistic expectations and poor decision-making. Various tools are used to conduct market analysis, including SWOT analysis (Benzaghta et al. 2021), PEST analysis (Tanya Sammut-Bonnici and David Galea 2014) and Porter's five forces analysis (De, Arons, and Waalewijn 1999). These tools help to assess the external environment by identifying opportunities and threats, market trends and competitive forces. SWOT analysis helps to identify a business's strengths, weaknesses, opportunities and threats, while PEST analysis examines the political, economic, social and technological factors that affect a business. Porter's five forces analysis assesses the competitive forces in the industry, including the bargaining power of suppliers, buyers and competitors. Conducting a market analysis provides the necessary information to create a successful business model. It helps determine the appropriate marketing, pricing, and product development strategies. In addition, it helps to identify potential risks such as changing market trends, new regulations or changes in consumer behaviour. (Benzaghta et al. 2021; Shabanova et al. 2015; Tanya Sammut-Bonnici and David Galea 2014; De, Arons, and Waalewijn 1999)

1.6.1 SWOT Analysis

The SWOT Matrix is a tool used for assessing businesses and their ability to reach their goals by evaluating their internal and external factors. It recognizes that certain features are within the business's control while others are outside. A combination of strengths, weaknesses, opportunities, and threats analyses can produce alternative options for a business.

Various methods have been combined with the SWOT analysis to provide accurate results in various contexts. A five-point scale item can improve SWOT findings and lead to high reliability. The AHP and ANP methods have been used to complete a SWOT analysis in a way that can bring insightful and accurate results. These methods consider tangible and intangible measures in decision-making and provide an easy decision-making technique. Incorporating a SWOT analysis

with other methods has produced consequential strategic decisions, such as using AHP in a SWOT model, which can be regarded as A'WOT, or the SWOT fuzzy ANP technique, which overcomes ambiguity and criteria effects for the distribution company. (Benzaghta et al. 2021; Agarwal et al. 2009)

1.6.1.1 TOWS matrix

The TOWS matrix builds on the foundations of SWOT analysis by providing a framework for developing alternative strategies based on logical combinations of these factors.

The TOWS matrix consists of four strategy groups: Strength-Opportunity (SO), Strength-Hreats (ST), Weaknesses-Opportunities (WO) and Weaknesses-Threats (WT). SO strategies use internal strengths to exploit external opportunities, while WO strategies aim to reduce internal weaknesses by exploiting external opportunities. ST strategies use strengths to avoid or reduce the effects of external threats, while WT strategies are defensive tactics aimed at reducing internal weaknesses and external threats. (Benzaghta et al. 2021; Agarwal et al. 2009)

The first step to preparing the TOWS matrix is to conduct a situational assessment using the SWOT technique. This involves collecting and analyzing data from various sources, such as employees, customers, and industry experts. The SWOT analysis results in a list of internal strengths, weaknesses, and external opportunities and threats.

The next step is to compare the critical factors with relevant standards or guidelines. For example, in the case of engineering faculty, stakeholder priorities can be compared to the points listed in the NBA manual.

Once critical factors are identified and verified, the TOWS matrix can be used to develop alternative strategies. Besides other things, the TOWS matrix involves brainstorming and discussion with key stakeholders such as department heads and senior faculty members. Combinations of SWOTs are provided for each strategy to produce a logical and understandable outcome.

Using the TOWS matrix, organizations can develop effective strategies that leverage their strengths, mitigate their weaknesses, and take advantage of external opportunities while minimizing the impact of external threats. It is a valuable strategic planning tool that enables organizations to plan for the future and achieve their goals. (Agarwal et al. 2009)

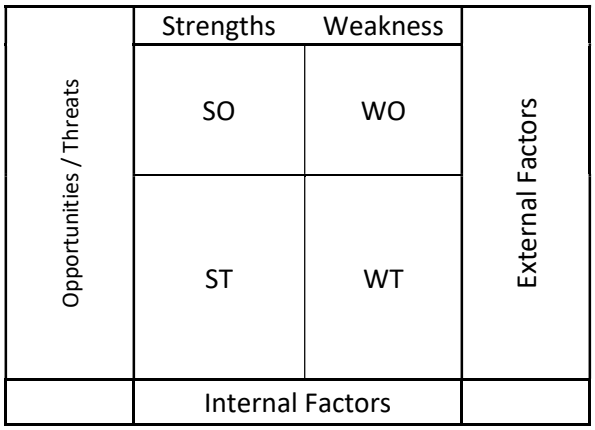


Figure 3 The SWOT Matrix (Benzaghta et al. 2021)

1.6.2 PEST Analysis

PEST analysis, also known as PESTLE analysis, is a strategic tool used to understand the impact of external macro-environmental factors on a firm's competitive position. The acronym PEST represents four sources of change - political, economic, social and technological. The external environment consists of variables that the firm cannot control but requires analysis for the firm's strategy to adapt to the changing business environment. PEST analysis is conducted for the whole firm, whether business units, relevant products, or a new venture or partnership.

Political factors refer to the extent to which politicians are likely to intervene in the commercial environment and are divided into supranational, national and subnational tiers. As commercial activities become more globalized, the supranational level becomes more critical. National policies influencing corporate strategy include fiscal policy, federal incentives for doing business, planning and permitting licensing procedures, and approval of new products and services.

The impacts of external factors are mitigated through a pre-emptive strategy, and opportunities are exploited because of new competitive positions that may be created in the process. Variants of the

PEST analysis tool are PESTEL or PESTLE (which adds environmental and legal components), STEEPLE (which takes a CSR perspective by adding an ethical dimension) and STEEPLED (which adds an analysis of demographic factors). The benefits of PEST and SWOT are further enhanced when used in conjunction with Porter's five forces model (including supplier externalities, consumer demand, new entrants, product substitution and competitive rivalry). (Tanya Sammut-Bonnici and David Galea 2014)

1.6.3 Porter's five forces model

Michael Porter's Competitive Forces Model, commonly known as Porter's Five Forces Model, is a widely used framework for evaluating profit potential in an industry. The model examines an industry's competitive environment and analyzes the collective strength of five fundamental forces that affect its profitability.

The five forces include:

The threat of new entrants: This force focuses on the ease with which new competitors can enter the industry. The threat of new entrants is high when it is easy for new firms to enter the market and low when there are significant barriers such as high startup costs, access to distribution channels, or government regulation.

Bargaining power of suppliers: This measures suppliers' control over the price and quality of inputs into the industry. The bargaining power of suppliers is high when they are few, when they provide unique inputs or when they can integrate and become competitors.

Bargaining power of buyers: This power measures the degree of control buyers have over the price and quality of industry outputs. Buyers' bargaining power is high when there are few buyers, and each buys large volumes when they can easily switch to other suppliers or are well-informed about the industry.

The threat of substitute products or services: This power considers the availability of substitutes that can fill the same need as the products or services of the industry. The threat of substitutes is

high when many substitute products or services are readily available, cheaper or more convenient, or offer better quality or performance.

The intensity of competitive rivalry: This strength measures the degree of competition between existing firms in the industry. The intensity of competitive rivalry is high when many competitors are roughly equal in size and strength and when the industry is growing slowly or not at all.
These five forces are based on structural components that impact the industry's profitability. The interplay of the five forces determines the sector's level of competition and overall profitability. Therefore, understanding these forces is crucial for strategic planning and decision-making by companies operating in the industry.

It is important to note that each force's potency may differ depending on the industry. A company's capacity to respond to each force will be determined by its different assets, skills, and competitive standing within the market. Porter's five forces model is used to analyze the profitability of an industry and predict how these factors will impact a firm's competitive position. (De, Arons, and Waalewijn 1999)

1.7 Design thinking

According to Shute, "analytic and creative process that engages a person in opportunities to experiment, create and prototype models, gather feedback, and redesign. " (Razzouk and Shute 2012)

Nowadays, "Design Thinking" is identified as an exciting new paradigm for dealing with problems in many professions, most notably Information Technology (IT) and Business. (Dorst 2011)

Design thinking is a problem-solving approach that combines analytical and creative processes to generate and test solutions. It is characterized by experimentation, prototyping, and gathering feedback to improve and redesign solutions. The key features of design thinking include visualization, creativity, and the ability to think outside the box. (Razzouk and Shute 2012)

The history of design thinking can be traced back to engineering, considered the primary or distinguishing activity. In the past, engineering curricula were primarily based on models focused on basic science, where students applied scientific principles to technological problems. However, this approach resulted in engineering graduates being perceived as needing help to practice in the industry. This concern led to recognizing the intellectual complexities and resources required to support good design education. In the 1980s and 1990s, design thinking began to receive increased attention in business settings as the design of products and services became a significant component of business competitiveness. Companies began to recognize the importance of design thinking to remain competitive in the market. (Razzouk and Shute 2012)

The term "design thinking" was coined by David Kelley, the founder of IDEO, a design and innovation consulting firm, in 1974 in his book The Universal Traveler: A Soft-Systems Guide to Creativity, Problem-Solving, and the Process of Reaching Goals. Since then, the term has been widely adopted and has become a popular method for solving complex problems in various fields, including business, education, healthcare, and government. (Camacho 2016)

1.7.1 Design thinking interview: In-depth interviews

In the book Design Thinking Research: Looking Further: Design Thinking Beyond Solution-Fixation (Leifer and Meinel 2019), Hasso Plattner discusses how design thinking can be used as a framework to foster innovation in products, services, and operations. He also highlights the need to deepen our understanding of how and why design thinking works, the factors that make it more successful than other approaches, and why it can fail. In the bold text, he mentions how the Design Thinking Research Program, a research initiative conducted jointly by himself Institute of Design at Stanford University in California and the Hasso Plattner Institute (HPI) for Digital Engineering in Potsdam, Germany. The program has been running for nearly ten years and has conducted over 100 studies examining, detailing, and making sense of design thinking in its many forms. Plattners emphasizes the importance of sharing this new knowledge and content with students, professionals, and experts to improve the teaching and learning of design thinking and make it accessible to all who seek to advance, drive, and support innovation and innovative culture for organizations or individuals making social and cultural change.

In later chapters, the book describes the aim to describe an online course about the design research phase and build a prototype MOOC for testing. The research method included conducting surveys to gather quantitative feedback and in-depth qualitative interviews with proto-MOOC learners after the course. The interviews were conducted with 16 participants. The interviewees were asked about their experience with the course content, learning modes, platform features, participant activity, and device choice. (Leifer and Meinel 2019)

The research process involves gathering feedback on a proto-MOOC (Massive Open Online Course) through user touchpoints. The researchers created the following measurement tools:

- pre-course evaluation surveys (CES)
- post-course evaluation surveys (CES)
- skill-confidence rating (SCR)
- qualitative interviews

The CES was used to compare learners' course expectations and satisfaction before the course started and after course completion, as well as gather data on basic demographics.

The SCR was used to assess the impact of the learning content on the learners' perception of their skill development. The SCR consisted of two pre and two post-ratings, which bracketed the Observation and Qualitative Interviewing modules. (Leifer and Meinel 2019)

1.7.2 Qualitative interviews

According to Shute, "analytic and creative process that engages a person in opportunities to experiment, create and prototype models, gather feedback, and redesign. " (Razzouk and Shute 2012) Nowadays, "Design Thinking" is identified as an exciting new paradigm for dealing with problems in many professions, most notably Information Technology (IT) and Business. (Dorst 2011)

Design thinking is a problem-solving approach that combines analytical and creative processes to generate and test solutions. It is characterized by experimentation, prototyping, and gathering

feedback to improve and redesign solutions. The key features of design thinking include visualization, creativity, and the ability to think outside the box. (Razzouk and Shute 2012)

The history of design thinking can be traced back to engineering, considered the primary or distinguishing activity. In the past, engineering curricula were primarily based on models focused on basic science, where students applied scientific principles to technological problems. However, this approach resulted in engineering graduates being perceived as needing help to practice in the industry. This concern led to recognizing the intellectual complexities and resources required to support good design education. In the 1980s and 1990s, design thinking began to receive increased attention in business settings as the design of products and services became a significant component of business competitiveness. Companies began to recognize the importance of design thinking to remain competitive in the market. (Razzouk and Shute 2012)

The concept of design thinking has evolved over the years and has been adapted to different fields and industries. The term "design thinking" was coined by David Kelley, the founder of IDEO, a design and innovation consulting firm, in 1974 in his book The Universal Traveler: A Soft-Systems Guide to Creativity, Problem-Solving, and the Process of Reaching Goals. Since then, the term has been widely adopted and has become a popular method for solving complex problems in various fields, including business, education, healthcare, and government. (Camacho 2016)

1.7.3 Design thinking interview: In-depth interviews

In the book Design Thinking Research: Looking Further: Design Thinking Beyond Solution-Fixation (Leifer and Meinel 2019), Hasso Plattner discusses how design thinking can be used as a framework to foster innovation in products, services, and operations. He also highlights the need to deepen our understanding of how and why design thinking works, the factors that make it more successful than other approaches, and why it can fail. In the bold text, he mentions how the Design Thinking Research Program, a research initiative conducted jointly by himself Institute of Design at Stanford University in California and the Hasso Plattner Institute (HPI) for Digital Engineering in Potsdam, Germany. The program has been running for nearly ten years and has conducted over 100 studies examining, detailing, and making sense of design thinking in its many forms. Plattners emphasizes the importance of sharing this new knowledge and content with students, professionals,

and experts to improve the teaching and learning of design thinking and make it accessible to all who seek to advance, drive, and support innovation and innovative culture for organizations or individuals making social and cultural change.

In later chapters, the book describes the aim to describe an online course about the design research phase and build a prototype MOOC for testing. The research method included conducting surveys to gather quantitative feedback and in-depth qualitative interviews with proto-MOOC learners after the course. The interviews were conducted with 16 participants. The interviewees were asked about their experience with the course content, learning modes, platform features, participant activity, and device choice. (Leifer and Meinel 2019)

The research process involves gathering feedback on a proto-MOOC (Massive Open Online Course) through user touchpoints. The researchers created the following measurement tools:

- pre-course evaluation surveys (CES)
- post-course evaluation surveys (CES)
- skill-confidence rating (SCR)
- qualitative interviews

The CES was used to compare learners' course expectations and satisfaction before the course started and after course completion, as well as gather data on basic demographics.
The SCR was used to assess the impact of the learning content on the learners' perception of their skill development. The SCR consisted of two pre and two post-ratings, which bracketed the Observation and Qualitative Interviewing modules. (Leifer and Meinel 2019) Daniel W. Turner III and Nicole Hagstrom-Schmidt describe Qualitative research design as a multifaceted endeavour that can be contingent upon the researcher's familiarity with a particular methodology. (Turner 2010) The interview protocol is more widely studied within the qualitative research design. In many cases, interviews are combined with other data collection forms to provide the researcher with comprehensive information for analysis. (Turner 2010)

As with quantitative analyses, various forms of interview design can be utilized to obtain rich, nuanced data utilizing a qualitative investigative perspective. Examples include:

- the informal conversational interview
- the general interview guide approach
- the standardized open-ended interview

Turner suggests that constructing practical research questions for the interview process is one of the most critical components of interview design. He references McNamara(McNamara 2009), who posits several recommendations for creating practical research questions for interviews, which include the following elements:

- The wording should be open-ended.
- Questions should be as neutral as possible.
- Questions should be asked one at a time.
- Questions should be worded clearly.

Turner then says that the final stage in the interview design process is interpreting the data gathered during the interview process. During this phase, the researcher must make sense of the information collected and classify the data into sections or groups of information, known as themes or codes. These themes or codes are consistent phrases, expressions, or ideas that were recurrent among research participants. (Turner 2010)

1.7.4 Interpreting qualitative interviews with design thinking

Karen von Schmieden, Lena Mayer, Mana Taheri and Christoph Meinel used the design thinking tool of feedback grids to structure notes from the 16 qualitative interviews they gathered in their research (von Schmieden et al. 2019). The detailed version was described in a table that presented the evaluation of the open post-skill confidence rating questions for two learning modules, Observation and Qualitative Interviewing. The feedback is organized into four categories:

Positive (+),
Negative (-),
Open Questions (?),
New Idea (!)

In the **Positive category**, the comments include feedback such as enjoyment, fun (to look out for daily workarounds), well-balanced content, and learning speed.

In the **Negative category**, the comments include feedback such as technical problems, unclear learning objectives/weekly goals, and the need for more knowledge on how to present a job story.

In the **Open Questions** category, the comments include similar feedback to the harmful category, such as the possibility of workarounds beyond products, unclear learning objectives/weekly goals, and a need for more knowledge on how to present a job story.

In the **New Idea** category, the comments include feedback such as enhancing the poignancy of topical session segments by adding more examples, more case studies, adding more content to observation skills, advice on interview setting, and receiving feedback on uploaded interviews.(Leifer and Meinel 2019; von Schmieden et al. 2019)

1.8 Data Governance

1.8.1 Data quality

Maximizing data quality is the aim of data quality management. Data quality management is a sub-function of data management, which comprises the planning, controlling, and provisioning of data assets. (Otto 2011) (Al-Ruithe, Benkhelifa, and Hameed 2019)

Low levels of data quality can have far-reaching consequences for a business, such as poor decision-making and missed business opportunities, since the data might not provide a clear picture of the circumstances. (Hooftman et al. 2018)Relying on manual data collection often causes a threat to the quality of data. Moreover, increased costs can occur when data have to be corrected at some point in the process. (Cichy and Rass 2019)A Gartner study claims that

poor data quality leads to an average loss of around $15 million for businesses. (Susan Moore, 2018) Especially in a Big Data environment, new challenges, costs and impacts arise, including the data's value, usability and overall quality. Tasks such as storage and processing of data but also the management of data quality, are critical in these environments. Poor data quality can also result in a compliance risk, i.e., when the standard of data quality does not match the expectations of supervisory authorities (Cichy and Rass 2019)

Otto (2011) defines data quality as something critical for enterprises to be able to meet a variety of business requirements. Examples of practices that enterprises should thus implement are; compliance with regulatory and legal provisions, integrated customer management ("360° view on the customer"), effective and efficient reporting ("single point of truth"), or integrated and automated business processes. (Otto 2011)

To show an example where data quality practice was required, Otto describes the situation of consumer goods manufacturer Nestle which is confronted with requirements from the French retail industry to provide "carbon footprint" information on the packaging of each product shipped to stores. This information must be specified as an attribute of the product data class. It has to be made available for the production and packaging process correctly, ultimately, and on time. (Otto 2011)

Such requirements particularly aggravate corporate data management (i.e. data that is used across the entire company) in large enterprises, which are typically characterized by complex and often globally spread organizational structures. Data Governance is a possible approach to meet these challenges as it specifies who makes decisions concerning specific data and the tasks and duties resulting from such decisions.

In the case of Nestle, Data Governance is applied to ensure that the right data source is used for providing correct information on carbon dioxide emissions and for specifying the time, form, and quality of this data, which is supposed to be available for the information being imprinted on the product labels. (Otto 2011)

1.8.2 Metadata

Jane Greenberg writes about metadata as data that describes other data, which is crucial for data science and analytics. (Greenberg 2017) According to him, metadata provides context and helps with the accurate interpretation and use of data by humans and machines. However, metadata research and development has yet to keep pace with big data and data science advances. Research literature and reports highlight the need for a principled approach to metadata management in significant data ecosystems and a need for research on metadata frameworks to ensure data trustworthiness. This gap raises questions about the disconnect between metadata and the larger sphere of data science research and how to address this challenge. (Greenberg 2017)

Different fields have different typologies for metadata, such as business and technical metadata in business and data warehousing or descriptive, technical, preservation, and usage metadata in digital libraries, archives, and repositories. These types of metadata are connected to the lifecycle of the object being represented or tracked. It is important to note that not all data with a meta status is labelled as metadata.

Examples include;

- Provenance Data
- Linked Data
- Contextual Data
- Authenticity data

These data exist because of the actuality of other objects and can only occur due to an object's activity. Metadata serves as an integrated layer in an information system, enabling the interplay between an object (data) and the desired activity (discovery) access and provenance tracking. (Greenberg 2017)

1.8.3 Data access

Data access refers to the user's ability to retrieve, view, and manipulate data stored in a database. Data's structural dependence and independence refer to the relationship between the database

schema (the structure and organization of the data) and the access programs used to retrieve and manipulate the data. (Rob, Coronel, and Crockett 2019)

Structural dependence means that a change in the database schema will affect data access, requiring changes in all access programs. On the other hand, structural independence means that changes in the database schema do not affect data access, and access programs can continue to be used without modification. (Rob, Coronel, and Crockett 2019)

Data dependence and independence also play a role in data access. Data dependence is a condition where data representation and manipulation depend on physical data storage characteristics. Changes in how data is stored will affect how it can be accessed and manipulated. Data independence, however, means that data access is unaffected by changes in the physical data storage characteristics.

The logical and physical data format can also impact data access. Logical data format refers to how a person views data within a problem domain, while physical data format refers to how a computer stores data. Ensuring data is stored in a logical format that aligns with the problem domain and is easily accessible to humans can make data access more efficient and effective. (Rob, Coronel, and Crockett 2019)

The most widely used query language for data access is Structured Query Language (SQL), which most DBMS (database management system) vendors support. SQL is a nonprocedural language, allowing users to specify what they want to do with the data without specifying how it should be done. Additionally, DBMSs provide APIs that allow access to the data through procedural languages such as COBOL, C, Java, Visual Basic.NET, and C#. These APIs allow developers to write programs that interact with the data stored in the database, making it possible to create, implement, monitor, and maintain the database. Additionally, the DBMS provides administrative utilities used by the database administrator and the database designer to manage and maintain the database, ensuring that data access and integrity are maintained. (Rob, Coronel, and Crockett 2019)

1.8.4 Data life cycle

Data life cycle management is critical for any enterprise or application where data is used and processed to produce results. Data availability for a certain period ensures accessibility and usability within the system. Data can be generated through different sources and available in various forms for accessibility. An extensive data-based application, such as the healthcare sector, generates a large amount of data through sensors and other electronic devices, which can be further classified into a model for report generation and predictions for various purposes for the benefit of patients and hospitals. (Rahul and Banyal 2020)

The data life cycle presents the entire data process in the system. The data lifecycle starts from creation, storage, usability, sharing, archiving and destruction in the system and applications. It defines the data flow within an organization. To successfully implement the model, there is a need to maintain the life cycle of data under a secured data management system. In their paper, Rahul Rohitash Kumar Banyal discusses the data life cycle with different steps and various works done for data management in different sectors and the benefits of the data life cycle for industrial and healthcare applications, including challenges, conclusions, and future scope. (Rob, Coronel, and Crockett 2019; Rahul and Banyal 2020)

1.8.4.1 Big data

Data is considered a valuable asset in industries today. Big data is an emerging subject under management information systems, computer science, and social science. The impact of big data reflects in social media, sensor systems used in various industries (including healthcare sectors), mobile devices, marketing, education, smart city, production, and e-commerce, among others. Big data is also widely suitable for IoT-based products and applications (Internet of Things). (Rahul and Banyal 2020; Greenberg 2017)

Big data analytics is used for applications where voluminous data is generated, which are highly scalable, stored at fault-tolerant distributed database systems, and processed effectively and efficiently. Big data is a term used for processing heterogeneous data (i.e. unstructured and structured data) for bringing it into an informative form. It passes through data extraction, transformation, and visualization to make the data effective for an application. Big data is collected,

stored, processed, and transformed big data into meaningful information through strategic tools and techniques. Big data is characterized by volume, variety, value, velocity, veracity, variability, visualization, etc. (Rahul and Banyal 2020; Greenberg 2017)

Big data analytics is a means to analyze and interpret any digital information and data. Technical and analytical advancements in big data analytics (BDA) determine the functional aspects of digital products and services. As with the fastest and growing data generations, it is essential to streamline and understand the process and mechanism of how big data analytics can add value to industries in several ways. Data life cycle management is essential in big data analytics, encompassing several steps. Big data analytics associated with different technical aspects comprise searching, mining, analysis, and usability, which is suitable and essential for enhancing any application. (Rahul and Banyal 2020; Greenberg 2017)

According to Rathul and Bayal, data life cycle management is essential to recognize and streamline to understand marketing strategies to popularize their product compared to other industries. Under data life cycle management, giant data sets exist, where it matters to examine which data suite is vital for whatever conditions. It reflects access conditions and significant data investment to produce underexplored business values. The corporate sector executive believes data life cycle management under extensive data engineering or big data analytics to be the new milestone to nurture business opportunities. They state that big data analytics is suitable for cost reduction, better and faster decision-making, and developing new products and services. (Rahul and Banyal 2020; Greenberg 2017)

During data acquisition, the degree of data consistency, completeness, benefits of external data usage, and benefits of unlimited data usage support in the decision-making process are vital factors to consider. The data acquisition process should maximize the benefits of external data usage while minimizing the risk of data inconsistency and incompleteness. (Rahul and Banyal 2020; Greenberg 2017)

To guarantee data consistency, it is essential to establish clear policies and procedures for data collection, management and storage. To ensure that data are accurate and complete, these activities

include defining data quality standards, establishing data management protocols, and implementing data validation procedures. The benefits of using external data in decision-making should also be considered. Statistics and context that may not be available from internal data sources can be obtained from external data. This may include consumer behaviour, market trends and competitor activity. (Rahul and Banyal 2020; Greenberg 2017)

2 Practical Part

2.1 The business model for the Data Governance Platform

A data governance platform is a software application that provides a framework for managing data assets within an organization. Companies must manage and control their data assets, ensure data quality and compliance, and minimize the risk of data breaches or leakage. Data governance platforms, typically used in large organizations with complex data structures and multiple data sources, are becoming increasingly necessary as companies collect and use more data.

To address this need, a data governance platform provides a unified approach to data management. Organizations can ensure that these policies are consistently enforced by establishing policies, procedures, and standards for data usage and quality. Implementing a data governance platform is particularly important because data has become a liability if it needs to be adequately managed and governed, leading to inefficiencies, errors, and risks.

The practical part of the thesis will focus on developing the business model for the new data governance platform. The practical part will involve conducting a market analysis to identify the platform's target market and assess the level of competition in the market. The practical part will also involve conducting a PEST analysis will evaluate the political, economic, social and technological factors that could influence the platform's success in the Czech Republic. In addition, a SWOT analysis will be conducted to identify the platform's strengths, weaknesses, opportunities and threats, which will help develop the TOWS matrix. The TOWS matrix will then be used to propose strategies for the platform. A Porter's analysis will also evaluate the market's competitive strength and identify the platform's competitive advantage. In addition, a unique value proposition for the platform will be proposed and described to differentiate it from its competitors.

The Lean Canvas model will be used to develop a business plan for the platform. Qualitative interviews will be conducted to gain insight into the needs and preferences of the platform's target market. A business case will be presented to introduce a more practical aspect of the model. Finally, a wireframe model will be developed to visualize the unique value proposition describing the platform's functionality and user experience.

2.1.1 Problem description

Over the years, large companies have gathered vast amounts of data; however, utilizing the information their data can provide often fails to be fully addressed. The new term “data-driven company” has recently grown in popularity, and the state of full data utilization is sought after. However, even when the companies decide to start utilizing their data, they are bound to face many obstacles, such as:

Data accuracy and quality

Data accuracy and quality have become a critical concern for organizations, as deviations from expected standards can have far-reaching consequences, including potential reputational damage and sub-optimal decision-making.

Outdated information

Another challenge is the persistent problem of outdated information. The problem with outdated information stems from a need for consistent data updates, leading to inconsistencies between different sources of information.

Data fragmentation

Data fragmentation across different systems and databases has also become a pervasive problem, presenting difficulties in accessing and analysing information. Data fragmentation can lead to duplicate and inconsistent data, undermining the reliability of a company's information systems.

Data Ownership

Data ownership and accountability remain challenging for many organizations as the division of responsibilities is often unclear and undefined. Data ownership can lead to efficiency, clarity, and effective data management and governance.

Big data management

The rapid expansion of big data has added a new level of complexity to the data management challenge. The scale of data has become immeasurable, requiring companies to have robust infrastructure and tools, as well as experimental processes to manage and understand the information.

2.1.2 Market Analysis

Data governance is integral to any organization's data management strategy to ensure accuracy, completeness, and data reliability. Organizations are constantly generating and collecting vast amounts of data in today's data-driven world, and effective data governance is becoming increasingly important.

As a result, the data management market has been expanding recently and is expected to continue its growth trajectory in the coming years. According to market research reports by Gartner (Ehtisham Zaidi 2019), the global data management market is expected to grow at a compound annual growth rate (CAGR) of 10.6% between 2020 and 2025.

This market growth can be attributed to several factors, including the increasing volume of data generated by organizations, the need for better data governance, and the growing demand for data management solutions that can help organizations ensure the quality and integrity of their data.
One of the significant challenges that organizations face in the data management market remains obsolete data. Outdated data can lead to incorrect decisions, reduce efficiency, and negatively impact an organization's reputation.

2.1.2.1 Existing solutions in the market

Several factors, including digital transformation, cloud-based solutions, and remote work, are propelling the Data Catalog Market at an impressive growth rate. Data catalogues are crucial tools for obtaining a consolidated perspective of data for enhancing decision-making due to businesses relying on massive amounts of data collected from numerous sources. The growth of datasets has increased the demand for data catalogues, promising new angles and insights to increase output and predict results.

Due to their many advantages, such as lower operational costs, simple deployments, and greater scalability in terms of networked resources, cloud-based data catalogue solutions are gaining significant traction. These solutions also increase working flexibility for businesses using real-time analysis and streamline real-time deployment. The lack of data management standards and enterprise concerns about data security and privacy hampers the market's expansion.

The Data Catalog Market is segmented according to the component, deployment mode, end-user industry vertical, and geography. The combined solution maximizes individual productivity, improves data quality, eliminates data duplication and makes data discovery easier. Due to the expanding use of data in numerous BI tools and the rising acceptance of digital technologies in industries like BFSI, healthcare, telecom and IT, and manufacturing, North America is anticipated to dominate the data catalogue market, which is expected to expand quickly throughout the forecast period.

With several significant players already present on a global scale, the Data Catalog Market is anticipated to become more fragmented. These market players implement cutting-edge strategies like partnerships, mergers, and acquisitions to gain a worldwide market share. Among the top players in the market are IBM, Informatica, Alteryx, Collibra, Alation, Microsoft, TIBCO Software, Datawatch Corporation, and Zaloni.

The market is highly competitive due to the rising demand for data catalogue solutions and services and the number of vendors providing these services. Vendors have been compelled by the competition to offer cutting-edge goods and services to satisfy the end-users expanding needs. Additionally, the market is seeing an increase in the amount of money that vendors are investing in R&D to bring new products and services to market that will meet the changing needs of end users.

IBM

IBM offers data management, enterprise content management, business analytics, and other IT solutions. The data catalogue solution from IBM has several features that let companies enhance their data governance and management procedures. A centralized metadata repository, automated data discovery and classification, and self-service analytics capabilities are just a few of the significant advantages of IBM's data catalogue solution. IBM is well-known worldwide and caters to clients in various sectors, including banking, energy, insurance, the automotive industry, and retail. (“IBM InfoSphere Information Governance Catalog | IBM” n.d.)

Informatica

Informatica is a significant player in the data catalogue market and provides various data management solutions. The solution offers several features, such as automated data profiling, data lineage tracking, and machine learning-based data discovery. With a strong emphasis on data quality, Informatica provides various tools to help companies ensure their data is reliable and consistent. (“Data Governance Tools and Products | Informatica” n.d.)

Alteryx

One of the top providers of data analytics solutions is Alteryx, which also offers a data catalogue solution that helps companies manage and analyze their data more efficiently. Data profiling, data lineage tracking, and data governance capabilities are just some features in Alteryx's data catalogue solution. The solution is very flexible and can be altered to suit the requirements of different businesses. With a strong emphasis on self-service analytics, Alteryx provides users with a wide range of tools to analyze and visualize their data. (“Alteryx Analytics Automation Platform | Alteryx” n.d.)

Collibra

Data governance business Collibra provides various data management solutions, including a data catalogue solution. With the help of Collibra's data catalogue solution, businesses can more efficiently find, comprehend, and manage their data assets. A central repository for metadata, data lineage tracking, and data quality monitoring are just a few features the solution has to offer. Collibra strongly emphasizes data governance and provides tools to support businesses in ensuring

their data complies with internal and external policies and regulations. (“Data Governance | Collibra” n.d.)

Alation

Alation is a data catalogue company that provides a cloud-based data catalogue solution made to aid businesses in finding, comprehending, and making better use of their data. The Alation solution has several features, such as automated data discovery, lineage tracking, and collaborative data governance capabilities. Since it can be tailored, the solution can be used by businesses of all sizes and is highly scalable. Alation strongly emphasizes collaboration and provides a wide range of tools to assist users in sharing and collaborating on data throughout the organization. (“Enterprise Data Catalog Software | Alation” n.d.)

2.1.3 T.D.G.P Platform description

This new data governance platform T.D.G.P offers an array of features incorporating the best elements of several other platforms. In addition to existing data management features such as collaboration tools, data pipeline management, data quality management, and data cataloguing, this platform includes an innovative new feature: a pipeline-based approach that automatically updates child documents or databases based on changes in a parent document or database. This feature streamlines workflows and increases the accuracy and quality of data management, setting it apart from competitors.

One of the main functions of the platform remains the data cataloguing capabilities. The platform collects and catalogues data assets from various sources and provides comprehensive metadata and contextual information about each purchase. Natural language search and browsing features simplify the search and discovery of data assets.

Data lineage capabilities are also a vital feature of this platform. Users can understand the data flow across the entire organization with information about the platform's end-to-end data pipeline. Those features enable organizations to make better data-driven decisions, manage data risks and ensure regulatory compliance.

The platform's data quality management features allow users to create rules and policies and automate data validation and testing. Users can track data quality metrics and collaborate on data quality issues, making it easier to ensure data reliability and accuracy.

The user interface of the platform is simple and intuitive. The platform's modern and straightforward interface makes finding and using data sources easy. It also provides a selection of dashboards and reports that offer insight into data assets, data usage and quality, making it easier for users to make informed decisions about data usage.

2.2 Unique value proposition

To address the identified needs on the market and gain a competitive advantage, the platform T.D.G.P could implement a feature involving an innovative data pipeline approach that automatically updates child documents or databases based on changes in the parent document or database. This feature sets it apart from the data management competition by streamlining workflows and increasing the accuracy and quality of data management.

To illustrate the potential benefits of this feature, we can use an example of a bank using the proposed data management platform. In this example, the bank's lending process involves several documents, including applications, credit reports and financial statements. With the dateline feature of the proposed platform, any changes to the credit report or financial statements in the loan application would automatically be reflected in all other loan-related documents, ensuring that all documents contain accurate and up-to-date information. This feature saves time, reduces the risk of error, and improves the bank's ability to make informed decisions and reduce risk.

Another example of how this feature could pose an advantage for another business could be car manufacturing; automatic updating of child documents when parent documents change within the platform could be used to streamline workflows and improve data management in the manufacturing process. When changes are made to an automotive part's original drawing or design, the platform could automatically update all related child documents to reflect the change, including the bill of materials, manufacturing instructions, and inventory records. This feature helps reduce errors and misunderstandings that could lead to production delays or quality control issues.

Several technologies could be used to implement the unique feature of the proposed data management platform, which consists of automatically updating data in child documents upon changes in parent documents. One possible way to do this is through already mentioned data lineage technology, which traces and displays the origin, movement and transformation of data, allowing the data lineage to be traced from its source to its final state. This technology could track relationships between parent and child documents and ensure that any changes to a parent document trigger an update in related child documents.

Another technology that could be used is data change capture (CDC), which captures and records all data changes in the database, allowing real-time updates to related data. This technology can be integrated with data line technology to update related child documents automatically when the parent document changes.

The proposed platform could also incorporate artificial intelligence (AI) and machine learning (ML) technologies to increase efficiency and effectiveness. For example, these technologies could predict which child documents will likely be affected by parent document changes and automatically update those documents. The platform could also use natural language processing (NLP) technologies to analyze the content of documents and identify relationships between them without requiring manual labelling by users or editors.

To build a data management platform, the software development team would have to follow a set of best practices to make the platform reliable, secure and scalable. One approach is to follow DevOps principles, emphasizing collaboration, automation, and continuous improvement throughout the software development lifecycle. DevOps practices allow the development team to work more efficiently and deliver high-quality software regularly.

To create a robust and scalable platform, the development team can use containerization technology, such as Docker, to package the application and its dependencies into a portable container that can run consistently across environments. Containerization can also ensure that an application runs smoothly and reliably, eliminating the need for developers to configure the

development environment manually. The team could also use infrastructure-as-code tools like Terraform to provision and manage the platform's infrastructure, including its servers, network, and storage resources.

To ensure the platform's security, the development team should follow best practices such as applying the principle of least privilege, implementing strong access controls, and regularly testing the platform for vulnerabilities. Additionally, the development team could use automated testing tools such as Jenkins to identify and address potential security issues early in development.

2.2.1 PEST Analysis

2.2.1.1 Political factors in the Czech Republic

The Czech Republic is a country that has undergone many political and economic changes in recent years. In 2019, the country performed well economically, with continued GDP growth, low public debt and stable macroeconomic conditions leading to positive international credit ratings. However, the COVID-19 pandemic in 2020 has significantly impacted the economy, politics, public health, and society. Despite government assistance, the country experienced a 5.6% decline in industrial production and a drop in demand.

The COVID-19 pandemic brought the Czech Republic's governance flaws, institutional framework, checks and balances, and civil society to light. Due to the pandemic's management by the prime minister, socioeconomic disparities and flaws in digital technology, education, and health were revealed. These flaws primarily impacted single mothers, low-income groups, and the urban middle class. The pandemic opened up possibilities for changing the nature of work and accelerating the adoption of robotic technology.

The recent Russian invasion of Ukraine has significantly affected the political and economic landscape in the whole European Union, including the Czech Republic. Sanctions against Russia and other actions to defend Ukraine's territorial integrity and sovereignty are part of the EU's response. Following the invasion, EU member states applied to join NATO, giving Ukraine candidate status. The EU consists of several institutions that work together, including the European

Council, the European Commission, the European Parliament, the Council of the European Union and the European Court of Justice.

The Czech presidential election of 2023 was pivotal for the country, with retired NATO general Petr Pavel winning 58.32 per cent of the vote. Pavel's background as a retired army general and his experience as chief of the Czech Army General Staff from 2012 to 2015 and chairman of the NATO Military Committee from 2015 to 2018 gave him the edge over his opponent Andrej Babiš, who won 41.67 per cent of the vote.

There is a contrast between Pavel's position and the previous two Czech presidents, Václav Klaus, a Eurosceptic, and Miloš Zeman, who was seen as a sympathizer of Russian President Vladimir Putin. Pavel sees the Czech Republic as culturally, economically and socially closer to the West and part of the world it traditionally belongs to. Pavel's "Stronger Together" initiative, which he launched in 2020 to help volunteer in the early days of the pandemic and help prepare the country for future crises, has also helped further cement his position as a strong and capable leader. ("EU - Political and Economic Environment" n.d.; Stiftung n.d.; "Petr Pavel Elected President of the Czech Republic - Prague, Czech Republic" n.d.)

2.2.1.2 Economic factors in the Czech Republic

In the upcoming years, the Czech economy will face several difficulties. The OECD predicts GDP growth will first slow from 2.4% in 2022 to -0.1% in 2023 before increasing to 2.4% in 2024. Rising energy prices, rising commodity prices, and disruptions in the gas supply are the leading causes of this slowdown. The price of living has also increased significantly due to Russian oil imports and the possibility of energy shortages. Activity in 2023 will be hampered by lower global growth, enduring constraints in global supply chains, and tight financial conditions.

Inflation will fall in 2023 but remain well above the 2% target. The unemployment rate will remain low, below 3%. The government has taken measures to counter the impact of high energy prices, including income support for households and small and medium-sized enterprises, energy price caps and a temporary reduction in fuel duty. At the same time, the central bank has tightened

monetary policy significantly to counter the rapid rise in core inflation and rising inflation expectations.

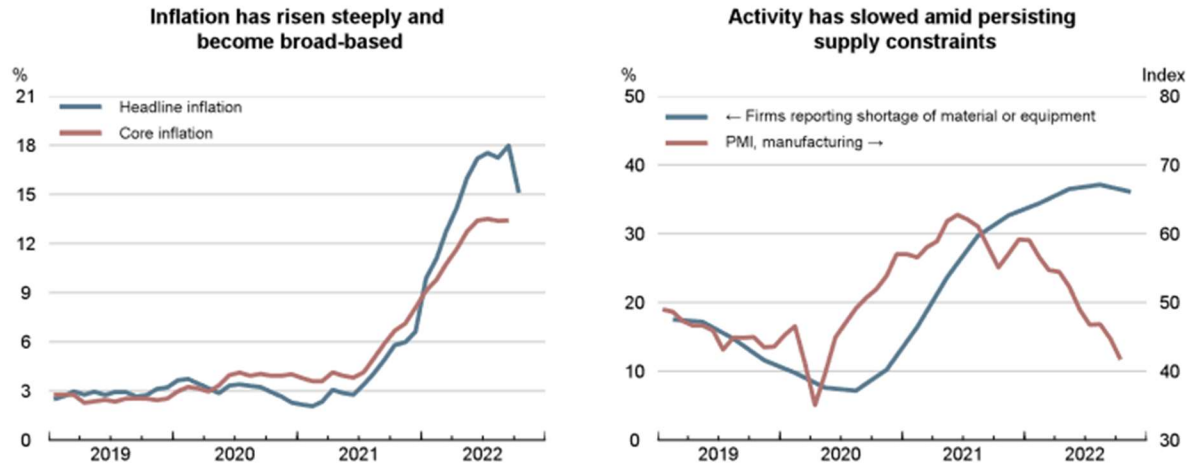


Figure 4 ("Czech Republic Projection Note OECD Economic Outlook November 2022 by OECD - Issuu" n.d.)

Before the war, the Czech Republic relied heavily on Russian energy imports, with almost all of its gas coming from that country. Efforts are being made to fill gas storage capacity and find alternate sources adequately. To lessen the effects of high energy prices, the government has implemented several policies, including price caps on energy, income support for households and small and medium-sized businesses, and a temporary reduction in fuel excise duty.

The aggressive conflict in Ukraine has made things worse for the world's supply chains and put more pressure on energy and commodity prices. Exports and activity have been constrained, particularly in the manufacturing and construction industries, due to supply disruptions of essential raw materials and components. Due to high levels of uncertainty and the impending energy crisis, real wages have fallen precipitously despite low unemployment, and consumer and business sentiments have deteriorated.

The cost of essential services and cash transfers increased due to the significant influx of Ukrainian refugees into the Czech Republic. By the end of August 2022, many working-age refugees had jobs, which helped ease labour market tensions and lessen reliance on government aid. The government must maintain a stringent and well-coordinated macroeconomic policy until inflation expectations are firmly in check. A steady consolidation of the structural fiscal balance should be

started to rebuild fiscal buffers and strengthen the labour supply. At the same time, the green transition should be accelerated to support growth and make it more sustainable. ("Czech Republic Projection Note OECD Economic Outlook November 2022 by OECD - Issuu" n.d.)

2.2.1.3 Socio-Cultural Factors in the Czech Republic

The Czech Republic is facing numerous social and cultural challenges as it tries to recover from the effects of the COVID-19 pandemic. The ongoing war in Ukraine is stifling its progress, and inflation and the cost of living are putting considerable social pressure on its citizens. Despite a tight labour market, vulnerabilities persist in the real estate market, leading to high uncertainty and risks to the economy and inflation.

Support for households and businesses is essential to address the cost-of-living crisis. However, measures that are targeted, temporary and maintain price signals to encourage energy conservation are preferable. Energy prices have risen significantly and are causing difficulties for vulnerable households. However, relief should be limited and temporary and interfere minimally with the price signals needed to encourage energy savings.

Active labour market policies should target sectors with significant shortages and provide training and up-skilling opportunities for less skilled workers to improve labour market integration. Linking retirement age to life expectancy and improving the participation of vulnerable groups would also benefit the labour market. Facilitating the inflow of migrants and effectively integrating Ukrainian refugees could help alleviate labour shortages.

Policies should promote technological innovation and worker reallocation, including investment in innovation strategies and significant reforms to the resilience and recovery plan. These policies would help build the technical and digital skills needed in high-productivity sectors and improve the business environment through streamlined regulation and simplified building permit processes.

Energy security depends on enhancing energy efficiency and expanding renewable energy supply through green transformation. It is possible to postpone the decommissioning of coal-fired power plants in the short term, but policies that promote more significant investment in renewable energy

should take precedence. Linking the retirement age to life expectancy is necessary for the pension system to be financially sustainable, as is increasing revenues through growth-promoting tax measures like higher environmental and property taxes and tighter compliance.
("Czech Republic: Staff Concluding Statement of the 2022 Article IV Mission", n.d.)

2.2.1.4 Technological Factors in the Czech Republic

The Czech Republic has made significant progress in developing Industry 4.0 technologies, as shown by the Czech Industry Analysis 2/2022 conducted by the National Centre for Industry 4.0. The analysis was based on 237 interviews with key representatives of selected companies in the Czech Republic and was conducted during the National Industry Summit held in June in Bethlehem Chapel. The analysis results show that nearly half of the country's manufacturing companies (49%) plan to invest more in new businesses in the next year, up an average of 15%.

Despite adverse economic conditions, with costs rising rapidly for manufacturing companies and 65% expecting a decline in profitability, the CEOs of these companies recognize the importance of investing in new technologies for their company's survival in a competitive global market. The COVID-19 pandemic has shown the importance of digital technologies to help companies overcome the impact of global crises by enabling remote communication with customers and deploying tools to monitor and optimize energy consumption, warehouses and production facilities.

Small and medium-sized enterprises (SMEs) in the Czech Republic are more likely to invest in new technologies than large companies, with 59% of SMEs planning to do so, compared to only 40% of large companies. However, the SME segment is also more volatile, with 27% of small businesses planning to reduce their investments by 20%. In contrast, large companies are more stable in their investments, with half of them (51%) spending the same amount on new technologies as in the first year.

Despite the technological advances of the last decade, the digital divide still needs to be solved in the Czech Republic. According to the Digital Economy and Society Index (DESI), the country ranks 19th among the 27 EU Member States and 15th regarding human capital, with 60% of Czechs having basic digital skills compared to the EU average of 54%. The Czech Republic faces

challenges in digital progress, such as the need for more school resources to ensure quality IT skills teaching and outdated approaches to teaching and curriculum development. In addition, the technology sector needs IT professionals, and many employers need help to train their existing workforce for current and emerging IT roles. This exacerbates the IT skills gap and makes it more difficult for communities to bridge the digital divide.

The digital divide is also a gender equality and equal opportunities issue in the Czech Republic. The country ranks low among European countries regarding gender balance in technology, with 90% of IT specialist positions held by men and a gender pay gap of more than 12%. Closing the digital gender gap requires quality ICT/IT training and education for women and girls. Czechitas, a leading organization that educates women about IT roles and connects them to job opportunities, aims to ensure that at least 30,000 women are employed in IT by 2025. The private and public sectors recognize the importance of building IT skills for women, as 9 out of 10 future jobs will require digital skills, and there will be nearly 1.7 million unfilled professional IT jobs by 2025. The Czech EU Presidency convened a Czech Digital Week in the first week of November 2022, which addressed digital education, the future of technology and the state of the online space, demonstrating the spirit of collaboration and the power of the Czech people to tackle these challenges. Investing in women and girls will determine what digital progress will be possible. ("Czech Industry Analysis 2022: Smart Industry 4.0 Technologies Help in the Crisis - News Service - Czech Technical University in Prague" n.d.; "Czech Republic Tackles Gender Digital Divide as Part of EU Presidency - Kafkadesk" n.d.)

2.2.2 SWOT Analysis

2.2.2.1 Strengths

The data governance platform T.D.G.P presents many strengths that position it favourably in the current technological environment in the Czech Republic. Firstly, it is characterized by a team of qualified and experienced developers and IT specialists, which serves as a basis for developing a product of high quality and reliability. The team's proven ability to create customized iterations of the platform based on specific client needs further highlights their ability to deliver customized solutions that meet the diverse requirements of today's business environments.

In addition, the platform benefits from close collaboration with academic experts whose knowledge and expertise enrich the development process with cutting-edge research, best practices and new perspectives. These academic links ensure the platform is built on a sound foundation and supports continuous innovation and improvement.

The team's past success with previous clients strengthens the new platform's position in securing funding from government grants and private loans. Securing these resources can provide a valuable boost to the platform's expansion, innovation and ability to meet the evolving demands of its clients continually.

Adding the platform's feature of automatically updating data in "child documents" when changes are made to the "parent document" could set it apart from competitors in the data management sector. This capability promises significant value to clients by streamlining workflow processes, reducing errors caused by outdated information and improving overall data quality and accuracy.

2.2.2.2 Weaknesses

While having several strengths, the data governance platform T.D.G.P may face some weaknesses that could affect its overall success. One potential area for improvement is the need for brand recognition and reputation, which could hinder the platform's ability to establish itself as a trusted and reliable solution in the competitive data management market. In addition, expanding the platform's reach and penetrating new markets may prove challenging, especially given the potential market saturation and competition it may face from established players.

Moreover, the platform's dependence on external funding and grant support could create vulnerabilities that may hinder its long-term sustainability. In addition, the complexity of the platform and its potential technical integration issues with clients' existing infrastructure could pose difficulties and cause delays or additional costs for clients.

Also, the platform's learning curve may make it challenging for clients to navigate effectively, discouraging adoption and requiring significant training and support investment to enable clients to use the platform to its full potential.

2.2.2.3 Opportunities

The data governance platform T.D.G.P faces several potential opportunities that can position it for success in the data management market. The growing demand for data management solutions is one of the most significant opportunities for a platform that can meet clients' needs by providing comprehensive, reliable, and innovative data management solutions. The platform can establish a strong market position and reputation by addressing this need.

The platform can also leverage strategic partnerships and collaborations with other companies in related industries to expand its reach and penetrate new markets. Such collaborations can provide new opportunities for distribution, cross-promotion and access to new customers. These collaborations allow the platform to establish itself in new markets and gain valuable market share.

Expanding the platform's operations into new geographic regions is another opportunity for growth. This strategy can increase the platform's customer base and provide an opportunity to establish a strong presence in new markets, thereby strengthening its market position and reputation.

The platform can further differentiate itself from competitors by developing new features and integrations to enhance its value proposition. The platform T.D.G.P can provide unique, tailored solutions that address each client's specific data management challenges by staying in tune with evolving client needs and investing in research and development.

The platform T.D.G.P can also leverage artificial intelligence and machine learning technologies to improve its accuracy, efficiency and effectiveness. AI would enable the platform to streamline its operations, improve its overall performance and position itself as an innovative leader in the data management space.

2.2.2.4 Threats

The proposed data governance platform T.D.G.P may face several potential threats that could impact its market competitiveness and relevance. One significant threat is the rapid pace of technological change, which requires the platform to be agile and adaptable to evolving market demands and emerging trends or risk becoming obsolete.

Another significant threat is cybersecurity risks, as data governance platforms handle sensitive data at risk of being hacked, stolen, or compromised. A security breach could result in severe financial and reputational damage to the platform and its clients.

The growing dominance of artificial intelligence (AI) and machine learning (ML) technologies represents another potential threat, as these technologies may eventually become the primary means of creating and managing data. This could reduce the need for human intervention and potentially undermine the value proposition of data governance platforms.

Geopolitical conflicts, such as the ongoing conflict between Ukraine and Russia, can result in economic uncertainty and instability, impacting the platform's financial performance and market position. This threat could require the platform to develop contingency plans to mitigate its potential impact.

It is also essential to consider that increasing competition in the data governance market poses a potential threat to the platform's market position. New players are regularly entering the market, and the platform must evolve and differentiate itself to remain competitive and relevant.

2.2.3 TOWS Matrix

The TOWS matrix is a strategic tool created based on a SWOT analysis for a new data management platform. The table outlines possible strategies that can be developed to minimize weaknesses and maximize opportunities. In particular, the matrix suggests leveraging internal strengths such as an experienced and skilled development team, collaboration with academic experts, and successful past projects to establish the platform's credibility and reputation. External opportunities, such as the growing demand for data management solutions, strategic partnerships, and collaborations, can be leveraged to mitigate dependence on external funding and grant support. On the other hand, external threats such as the rapid pace of technological change, cybersecurity risks, and the dominance of AI and ML technologies can be addressed by leveraging the platform's internal strengths and developing effective strategies to compete with other companies in the market.

TOWS Matrix	Internal strengths	Internal weaknesses
	The experienced and qualified development team	Lack of brand recognition
	Collaboration with academic experts	Dependence on external funding
	Successful past projects	The platform's complexity
External Opportunities	Successful past projects can be highlighted in commercial or academic articles to establish the platform's credibility and reputation.	Utilize strategic partnerships and collaborations to mitigate dependence on external funding and grant support in various competitions. Ask for recommendations or letters of interest.
Growing demand for data management solutions		
Strategic partnerships and collaborations	Strategic partnerships and collaborations with other companies in related industries or universities, such as CZU VŠE or ČVUT.	Use customer feedback to develop a comprehensive and user-friendly knowledge base. Create articles based on the most common questions. Allow users to ask the user community and publish responses by the community after validating them by platform administrators.
Expansion into new regions		
External Threats	Utilize the development team and academic experts to research and develop AI and ML technologies. Educate the team and support a knowledge-sharing culture to promote growing expertise.	Increase marketing efforts to raise awareness of the platform and its value proposition to potential clients, attend public events and organize events.
The rapid pace of technological change		
Cybersecurity risks	Utilize successful past projects as evidence of the platform's ability to integrate new technologies, learn from mistakes and findings from past projects, and create an internal knowledge database with those findings.	Collaborate with academic experts to incorporate best practices for the platform's development leverage the AI and ML technologies
The dominance of AI and ML technologies		

Table 1TOWS Matrix

2.2.4 Porter's analysis

Porter's five forces model provides a valuable framework for analyzing the competitive landscape of a data management platform.

2.2.4.1 The threat of new entrants

The threat of new entrants is relatively high, given the low entry barriers and the growing data management solutions market. For example, data management software company Collibra went public in 2021 with a market capitalization of over $5 billion, doubling its worth from 2020, indicating significant growth potential in the industry. ("Collibra Doubles Valuation as Investors Chase next Cloud Winner", n.d.) However, established players such as IBM, Informatica and Oracle may discourage new entrants due to their significant market share and brand recognition. ("Data Catalog Market Analysis - Industry Report - Trends, Size & Share", n.d.)

2.2.4.2 The bargaining power of vendors

The bargaining power of vendors is relatively low because several technology vendors can provide the necessary tools and infrastructure for data management platforms.

2.2.4.3 Bargaining power of buyers

The bargaining power of buyers is high due to the increasing number of data management solutions available in the market, leading to price sensitivity and intense competition.

2.2.4.4 The threat of substitute products or services

The threat of substitute products or services is relatively low because data management platforms are essential for companies to comply with regulations, manage data quality, and reduce operational risks. Emerging technologies such as artificial intelligence and machine learning can still disrupt the market and reduce the need for traditional data management solutions. ("What Is Machine Learning Data Catalogs? Benefits and Use-Cases", n.d.)

2.2.4.5 The intensity of competitive rivalry

The intensity of competitive rivalry is high due to the many players in the market and the constant need for innovation and differentiation. Established players like IBM, Informatica and Oracle have

significant market share and brand recognition. ("Data Catalog Market Analysis - Industry Report - Trends, Size & Share", n.d.)However, smaller players such as Collibra, Talend and Alation are snowballing and are challenging the incumbents. . (“Data Governance Market Analysis, Size And Trends Global Forecast To 2022-2030” n.d.)

2.2.4.6 The threat of external factors

The threat of external factors such as geopolitical conflicts, the rapid pace of technological change, and cybersecurity risks pose significant challenges to the growth and success of data management platforms. For example, the recent hacking of SolarWinds (“How Russia Used SolarWinds To Hack Microsoft, Intel, Pentagon, Other Networks : NPR” n.d.) highlights the significant risks and vulnerabilities that can be associated with data management platforms, requiring the platform to be agile and adaptable to new trends and threats.

2.2.5 Lean canvas

The Lean Canvas created for the Data Governance T.D.G.P platform is inherently connected to PESTLE, SWOT and Porter analyses. Conducting these analyses was essential to capture the various factors that could influence the success and longevity of the platform. By examining political, economic, sociocultural, technological, legal and environmental factors in a PESTLE analysis, the author have identified key trends and external elements that can influence the platform's growth, position and opportunities. The SWOT analysis yielded valuable insights into the platform's internal strengths and weaknesses and external opportunities and threats. By recognizing these aspects, the author have ensured that the Lean Canvas platform effectively leverages its strengths, addresses its weaknesses, and capitalizes on opportunities in the marketplace. For example, the disadvantages section of Lean Canvas emphasizes the platform's experienced development team, successful previous projects and customized solutions identified as strengths in the SWOT analysis. Porter's analysis helped us understand the competitive landscape in the data management market, which informed the Existing Alternatives and Key Metrics sections of the Lean Canvas. Recognizing the fierce competition in the market, the author have emphasized the need for effective marketing and brand recognition efforts in the Channels and Cost Structure sections. The author also highlighted the importance of customer support, including training and user support, to help clients overcome potential learning curves and navigate the platform effectively.

2.2.6 Lean canvas template

Lean Canvas Data Governance platform T.D.G.P

Problem	Solution	Unique Value Proposition	Unfair Advantage	Customer Segments
Data accuracy and quality Outdated information Data fragmentation Data Ownership Big data management	Data cataloguing capabilities Data line capabilities Data quality management User interface	Our platform's innovative data pipeline approach automatically updates child documents, streamlining workflows and improving data accuracy. AI and machine learning enhance efficiency—a reliable, secure, and scalable solution offering significant benefits to multiple industries.	Cooperation with academic institutions via connections at the grounds of CZU + Experienced and qualified development team, Successful past projects Customizable platform	Enterprise businesses Small and medium-sized businesses (SMBs) Data scientists and analysts Government institutions
Existing Alternatives	**Key Metrics**	**High-Level Concept**	**Channels**	**Early Adopters**
Collibra Informatica Alation Talend IBM InfoSphere Information	Adoption rate User satisfaction Data quality Compliance Time-to-value: Return on investment (ROI) Platform usage	Data governance platform = "GitHub for data" Just as GitHub provides a platform for developers to collaborate on and manage code, a data governance platform can provide a similar platform for data professionals to collaborate on and manage data.	Creating long-form content, such as comprehensive guides or reports Utilizing LinkedIn's targeting capabilities Attending industry conferences Sending regular newsletters Running targeted search ads Emphasized the importance of marketing and brand recognition efforts	Organizations with complex data environments Data-driven organizations Organizations with significant data growth Organizations seeking customizable solutions
Cost Structure		**Revenue Structure**		
Research and development costs		Subscription fees		
Infrastructure costs		Data quality certification fees		
Marketing and sales costs		Professional services fees		
Customer support costs				
General and administrative costs				

Table 2Lean canvas model

2.3 Project Schedule and Time Analysis

2.3.1 Definition of project activities and team roles

Project Initiation

Roles: Project Manager, Business Analyst, IT Manager

The project scope and objectives will be defined during this phase, and the team will be assembled. The project timelines and milestones will be established, and the initial project plan and budget will be developed.

Requirements Gathering

Roles: Project Manager, Business Analyst, IT Manager, UX Designer, Developer

This phase involves identifying key stakeholders, gathering requirements, conducting market research and analysis, developing use cases and user stories, and defining project specifications and deliverables.

Design and Architecture

Roles: Project Manager, IT Manager, UX Designer, Developer

In this phase, the platform architecture and infrastructure will be developed. The user interface and user experience will be designed, the database schema and data models will be developed, and the system requirements and design will be defined.

Development and Testing

Roles: Project Manager, IT Manager, Developer, QA Engineer

The platform features and functionality will be implemented in this phase, and unit, integration, and system testing will be conducted. User acceptance testing will be performed, and documentation and training materials will be developed.

Development Phase:

Roles: Project Manager, IT Manager, Developer, QA Engineer

During this phase,

- requirements gathering and analysis
- architecture and design
- database design and development
- backend development
- frontend development
- integration and integration and testing

Testing Phase: Roles: Project Manager, IT Manager, QA Engineer

The testing phase will include unit testing

- integration testing
- system testing
- acceptance testing
- performance testing
- security testing
- and user acceptance testing

Post-Deployment Support

Roles: Project Manager, IT Manager, Developer, QA Engineer, Customer Support Specialist

This phase involves providing user training and support, monitoring system performance and stability, conducting bug fixing and software maintenance, collecting user feedback and implementing improvements.

Roles for the project:

Project Manager - 1 employee

Business Analyst - 1 employee

IT Manager - 1 employee

UX Designer - 1 employee

Developer - 12 employees

QA Engineer - 3 employees

Customer Support Specialist - 1 employee.

2.3.2 Determination of activity sequences

1. Define project scope and objectives
2. Assemble project team
3. Establish project timelines and milestones
4. Develop initial project plan and budget
5. Requirements Gathering
6. Design and Architecture
7. Development and Testing
8. Development Phase:
    1. Requirements gathering and analysis
    2. Architecture and design
    3. Database design and development
    4. Backend development
    5. Frontend development
    6. Integration and testing
9. Testing Phase:
    1. Unit testing
    2. Integration testing
    3. System testing
    4. Performance testing
    5. Security testing
10. Post-Deployment Support

WBS ID	Activity Description	Predecessor Activity
1	Project Initiation	None
1.1	Define project scope and objectives	1
1.2	Assemble project team	1.4
1.3	Establish project timelines and milestones	1
1.4	Develop initial project plan and budget	1.1,1.3
2	Requirements Gathering	1.4
2.1	Identify key stakeholders and gather requirements	1
2.2	Conduct market research and analysis	1
2.3	Develop use cases and user stories	2.1
2.4	Define project specifications and deliverables	1.2, 2.2
3	Design and Architecture	2.4
3.1	Define system requirements and system design	2.4
3.2	Develop platform architecture and infrastructure	3.1,3.3
3.3	Design user interface and user experience	2.3
3.4	Develop database schema and data models	3.2
4	Development and Testing	3.4
4.1	Implement platform features and functionality	3.4
4.2	Conduct unit, integration, and system testing	4.1
4.3	Perform user acceptance testing	4.1
4.4	Develop documentation and training materials	4.3,4.2
5	Post-Deployment Support	4.4
5.1	Provide user training and support	4.4
5.2	Monitor system performance and stability	4.4
5.3	Conduct bug fixing and software maintenance	4.4
5.4	Collect user feedback and implement improvements	4.4
6	Project Completion (End of Project)	5.1, 5.2, 5.3, 5.4

Table 3 WBS Project Schedule and Time Analysis

2.3.3 PERT method

In this diploma thesis, the PERT (Program Evaluation and Review Technique) method was chosen over the Gantt chart for time analysis due to its several key advantages that will benefit the time analysis of the proposed development of the data governance platform. The PERT method is particularly well-suited for projects with high uncertainty and complexity, as it allows for a more comprehensive understanding of the potential variability in task durations. Using probabilistic time estimates for each activity, PERT provides a more realistic representation of the project timeline, considering the inherent uncertainties that may arise during the project's execution.

On the other hand, Gantt charts rely on deterministic time estimates, which may not accurately capture the variability and uncertainties in project tasks. This can lead to overly optimistic or pessimistic projections, ultimately hindering effective decision-making and resource allocation. Furthermore, Gantt charts are primarily focused on presenting a visual representation of the project schedule, lacking the analytical capabilities offered by the PERT method.

Additionally, the PERT method emphasizes the identification of critical paths within the project, which are sequences of tasks that directly impact the project's overall duration. By highlighting these critical paths, the PERT method enables project managers to focus their efforts on the most crucial activities, ensuring that resources are allocated efficiently, and potential bottlenecks are addressed in a timely manner.

Stage	Activity ID	Optimistic	Most Likely	Pessimistic	Duration (Te)	Standard deviation (σ)	Variance (σ^2)	Predecessor Activity/ies
Project Initiation	1							None
	1.1	1	2	3	2	0.5	0.25	1
	1.2	1	4	7	4	1	1	1.4
	1.3	2	3	4	3	0.33	0.11	1
	1.4	2	4	6	4	0.67	0.44	1.3, 1.1
Requirements Gathering	2.1	1	3	5	3	1.33	1.78	1.4
	2.2	2	4	6	4	0.67	0.44	1, 2.4
	2.3	2	3	4	3	0.33	0.11	2.1
	2.4	2	4	6	4	0.67	0.44	1.2, 2.2
Design and Architecture	3.1	3	6	9	6	1.33	1.78	2.4
	3.2	4	7	10	7	1.67	2.78	3.1, 3.3
	3.3	2	5	8	5	1.33	1.78	2.3
	3.4	3	6	9	6	1.33	1.78	3.2
Development and Testing	4.1	5	8	11	8	1.33	1.78	3.4
	4.2	2	5	8	5	1.33	1.78	4.1
	4.3	2	4	6	4	0.67	0.44	4.1
	4.4	3	5	7	5	0.67	0.44	4.2, 4.3
Post-Deployment Support	5.1	3	5	7	5	0.67	0.44	4.4
	5.2	3	6	9	6	1.33	1.78	4.4
	5.3	3	5	7	5	0.67	0.44	4.4
	5.4	2	4	6	4	0.67	0.44	4.4
End of project	6							5.1, 5.2, 5.3, 5.4

Table 4 Project phases Project Schedule and Time Analysis

2.3.4 Project paths

Critical Path: Follow the longest-duration path, considering the predecessors.

Path: 1.1 → 1.4 → 1.2 → 2.4 → 3.1 → 3.2 → 3.4 → 4.1 → 4.2 → 4.4 → 5.2 → 6

Total Duration (Te): 2 + 4 + 4 + 6 + 7 + 6 + 8 + 5 + 5 + 6 + 4.33 = 57.33

For the optimistic, realistic, and pessimistic paths, we'll consider the predecessor constraints and find the paths with the lowest duration for each case:

Optimistic Path: minimum duration for each stage, considering the predecessors.

Path: 1.1 → 1.4 → 1.2 → 2.4 → 3.1 → 3.3 → 3.2 → 3.4 → 4.1 → 4.3 → 4.4 → 5.1 → 6

Total Optimistic Duration: 1 + 2 + 1 + 4 + 3 + 2 + 4 + 5 + 2 + 3 + 3 = 30 weeks

Realistic Path: the most likely duration for each stage, considering the predecessors.

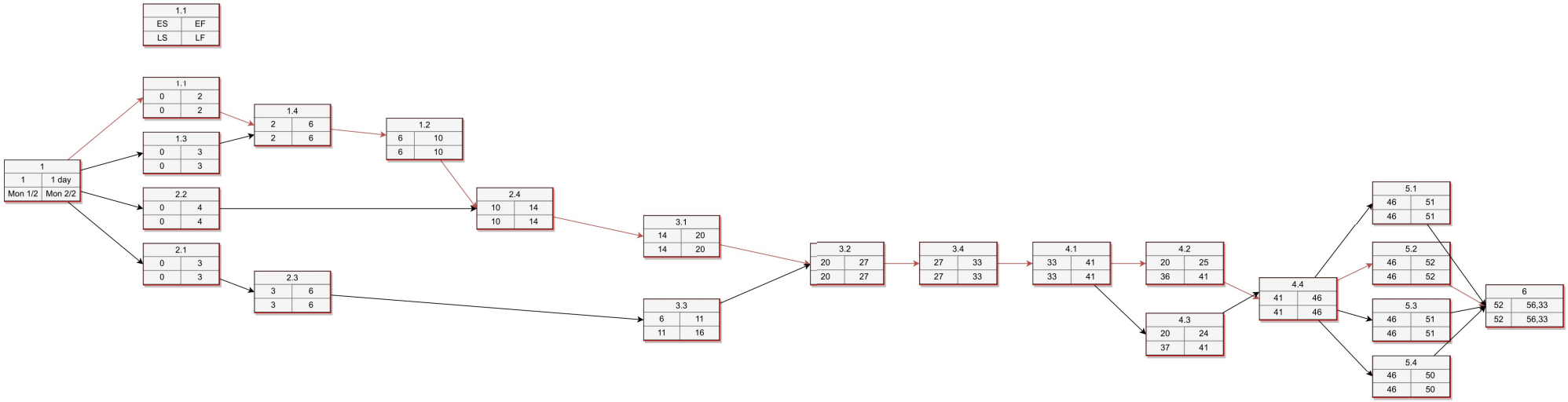
Path: 1.1 → 1.4 → 1.2 → 2.4 → 3.1 → 3.3 → 3.2 → 3.4 → 4.1 → 4.3 → 4.4 → 5.1 → 6

Total Most Likely Duration: 2 + 4 + 4 + 4 + 6 + 5 + 7 + 6 + 8 + 4 + 5 = 59 weeks

Pessimistic Path: the maximum duration for each stage, considering the predecessors.

Path: 1.1 → 1.4 → 1.2 → 2.4 → 3.1 → 3.3 → 3.2 → 3.4 → 4.1 → 4.2 → 4.4 → 5.2 → 6

Total Pessimistic Duration: 3 + 6 + 7 + 6 + 9 + 8 + 10 + 9 + 11 + 8 + 7 = 84 week



Picture 1PERT Project paths Data governance platform development with critical path highlighted.

3 Buyers' personas

Developing a new data governance platform requires a deep understanding of the target audience and their unique needs and challenges.

Data governance is a complex and multi-dimensional area that requires understanding the various stakeholders' requirements. For example, the requirements of a data analyst are very different from those of a project manager, and a technical manager's requirements differ from those of a business user.

The author has carefully defined and established this thesis's buyer personas. These personas embody the stakeholders involved in data management and governance.

- business user
- data analyst
- decision-maker
- technical lead

A comprehensive definition of these personas was achieved through market research and customer data analysis of potential competitors, which provided a thorough understanding of each persona's requirements, motivations, and challenges. These personas represent the diverse perspectives and requirements of those interacting with data within the organization and provide valuable insight into the complexities of data governance and management.

By creating a set of buyer personas for the various stakeholders involved in data governance and management, the author gained valuable insights into the different perspectives, requirements and challenges of those interacting with data within the organization. Creating personas helped the author understand the different types of buyers that may be encountered in the marketplace.

- Positive personas represent ideal customers who are knowledgeable and proactive in their approach to data management. These individuals clearly understand their requirements and challenges and are eager to adopt a platform that can help them achieve their goals. They are motivated by the need to effectively manage and protect their data and leverage it for business growth and success.

- Neutral personas, conversely, embody customers with limited data management knowledge and are still determining the best solution for their needs. They may need to familiarize themselves with the various data management platforms and evaluate their options.

- Negative personas represent customers with a negative perception of data management and may have previously experienced issues with data management solutions. They may be sceptical of new platforms and require additional persuasion and reassurance to adopt a data management platform.

Business User			
Positive /Neutral/ Negative	Positive	Neutral	Negative
Name	Amanda Smith	Nicol Natal	John Doe
Age	35	31	40
Occupation	Marketing Manager	Business Manager	Operations Manager
Company	A mid-sized B2B software company	A mid-sized retail company	A small retail company
Personality	Driven and highly organized, with a strong focus on results. She is a quick learner and constantly seeks to streamline her work processes. Open-minded and appreciates new technologies that can help her achieve her goals.	Practical and results-driven individual focused on achieving the best outcomes for his company. Not particularly passionate about data or technology. Generally sceptical of new technology and processes but is willing to try them.	Sceptical and resistant to change, preferring to stick with familiar ways of doing things. Often busy and needs more time to invest in learning new technologies. Easily frustrated and prefers to avoid dealing with complex systems.
Goals & Challenges:	She wants to make data-driven decisions that lead to better outcomes for her company. She needs help finding relevant data quickly and often spends too much time searching for the needed information . She wants to ensure that her team uses up-to-date data to inform their decisions.	He wants access to the data he needs to make informed business decisions but is not particularly interested in how it is collected or managed. They are frustrated by the current time-consuming and manual data collection and validation processes and the risk of using outdated or inaccurate data. He is still determining if a data governance platform would be a good investment for his company.	He wants to keep his work processes as simple as possible and does not need a data governance platform. He does not trust the accuracy of the data he has been provided and is wary of relying on it for decision-making. It needs more resources to allocate time and energy towards implementing a new system.
How a Data Governance Platform would help:	A data governance platform that provides a centralized repository for all data would make it easier for Amanda to access the information she needs. Having a clear understanding of the lineage of the data would give Amanda the confidence to rely on the information she uses. A platform that helps her manage and track data usage across her team would ensure everyone uses the most accurate information.	Appreciate the easy access to accurate and up-to-date data provided by a data governance platform. The platform's data lineage and version control features would give him confidence in the data, reducing the risk of using outdated or inaccurate data. See concrete evidence of the platform's benefits and return on investment before committing to its use.	He is sceptical about the value a data governance platform would bring to his work and does not see the need for it. He was worried that using a platform would introduce complexity to his work process and be difficult to understand and use. He does not want to invest time and resources into implementing a new system that may not deliver the results he is looking for.

Table 5 Business User Personas

Data Analyst			
Positive /Neutral/ Negative	Positive	Neutral	Negative
Name	Sarah Smith	Igoir Novak	Šárka Doležalová
Age	29	44	49
Occupation	Junior Data Analyst	Data Analyst	Senior Data Analyst
Company	large financial services company	large financial services company	Medium-sized IT company
Personality	Highly analytical and detail-oriented, and passionate about working with data. Problem solver and enjoys finding creative solutions to complex challenges. Proactive and has a solid drive to improve processes continuously and work more efficiently.	Detail-oriented and analytical with a solid technical background in data and statistics. Interested in the technical aspects of data management and governance but needs a strong business perspective.	Highly skilled data analyst with over ten years of experience Analytical and detail-oriented Experienced with various data analysis tools and techniques Not interceded in changes, thinks their approach is sufficient.
Goals & Challenges:	She wants easy access to accurate, up-to-date data to make informed decisions and support her company's business goals. Frustrated with manual and time-consuming data collection and validation processes. She is concerned about using outdated or inaccurate data in her analysis.	He wants access to clean and accurate data that will allow him to analyze and make decisions. Frustrated with the need for more standardization and consistency in data naming conventions, definitions, and formats. Unsure how a data governance platform would fit into his current workflow and data management tools.	Quickly and effectively analyze data to support business decisions Avoid dealing with outdated data Minimize time spent cleaning and verifying data Improve collaboration among team members Effectively communicate findings to stakeholders
How a Data Governance Platform would help:	Benefit from a data governance platform with a centralized and organized data repository. The platform would help her quickly retrieve and validate the data she needs, reducing the time and effort required for manual data collection and validation processes. The platform's data lineage and version control features also give it the confidence to use the data in its analysis, reducing the risk of using outdated or inaccurate data.	Appreciate standardization and consistency of data naming conventions, definitions, and formats provided by a data management platform.	Centralized data management system for better data organization and accessibility Automated data verification and cleaning to reduce time spent on manual tasks Real-time data updates to ensure data is always accurate and up-to-date Improved collaboration through streamlined project management and communication tools Increased confidence in data and insights due to better data governance and control.

Table 6 Data Analyst Personas

Decision Maker			
Positive /Neutral/ Negative	Positive	Neutral	Negative
Name	Aether Manel	Ara Diggory	Davidu Mattias
Age	44	45	31
Occupation	Director of Operations	Director of Data Management	CIO
Company	Large-scale Manufacturing Company	Medium-sized Manufacturing Company	Struggling Tech Firm
Personality	Experienced leader with a track record of making data-driven decisions Confident and assertive with a solid ability to make decisions under pressure Strategic thinker who always considers the big picture and long-term goals	data-driven professional with a balanced view of data management values data management accuracy and efficiency but also recognizes compliance and security's importance takes a considered decision-making approach, weighing each decision's facts and potential impacts.	Sceptical, Cynical, Demanding
Goals & Challenges:	Make informed decisions based on accurate and up-to-date data Ensure the company stays ahead of the competition Improve operational efficiency and productivity Maintain data security and privacy	balance accuracy and efficiency in data management while ensuring compliance and security. faces challenges in implementing data governance solutions that meet the needs of all stakeholders and are flexible enough to adapt to changing regulatory requirements.	Minimize costs while still meeting company goals and complying with regulations To find a data management solution that is easy to implement and maintain Trying to convince senior management to invest in a new solution
How a Data Governance Platform Would help:	Access to real-time data updates and insights Improved data quality and accuracy through automated data verification and cleaning processes Centralized data management system for better organization and accessibility Improved decision-making through real-time data analysis and visualization Increased confidence in data-driven decisions and insights due to better data governance and control.	Streamline the data governance process and ensure compliance with regulatory requirements. Could monitor data quality, track data lineage, and enforce data policies.	Demonstrate the cost savings and efficiency improvements that a data governance platform can bring Show its ease of use and maintenance compared to other solutions Demonstrate its ability to meet the company's compliance requirements and reduce the risk of financial penalties and legal action.

Table 7Decision Maker Personas

Tech Lead			
Positive /Neutral/ Negative	Positive	Neutral	Negative
Name	Maeva Uzochi	Tim Johnson	Maria Lin
Age	36	32	40
Occupation	Technical Lead	Technical Lead	Technical Lead
Company	Small SW company	Agriculture Company	medium-sized software company
Personality	Results-driven, hands-on, proactive, and strategic thinker,	a careful and practical individual who values stability and reliability in his work. He is meticulous in his approach and always prioritizes accuracy over speed. He has a solid technical background and stays abreast of the latest technological advances in his field.	a Demanding and highly sceptical individual who questions the value of technological goods and services. She approaches her work with an organization, frequently assessing products based on their features and capabilities.
Goals & Challenges:	Ensure that the company's IT infrastructure and systems operate efficiently and effectively to support business objectives To stay abreast of the latest technologies and industry trends to drive innovation within the company To find ways to reduce technical debt and streamline processes for better productivity	Ensure that the systems and processes in his organization run smoothly and efficiently. He also strives to maintain high data quality standards while balancing the needs of various departments within the organization. The biggest challenge is implementing new technologies that meet the organization's evolving needs while ensuring that current systems remain stable.	Interested in new products and technology solutions that claim to address the data management and governance issues her company is currently experiencing. She is looking for a solution that offers robust data protection features because she is very concerned about security and privacy issues related to data. Needs to spend more time on technical problems or employee training, searching for a user-friendly and simple solution.
How a Data Governance The platform would help:	Ensure that data within the company is appropriately managed and governed, reducing the risk of data breaches and compliance issues. It would provide a centralized system for managing and tracking data, making it easier to monitor the status of company data and address any issues.	Allow him to control the entire data lifecycle while ensuring that the data remains accurate and consistent. The platform would provide the tools to enforce data quality standards and manage data access rights, ensuring the organization's data remains safe and secure.	Not interested in a data governance platform that demands robust security features. Due to her limited time and financial resources, she is also unlikely to choose a problematic platform or require high technical knowledge. Maria is a picky client who needs a highly functional platform that meets her strict standards and specifications.

Table 8 Tech Lead Persona

3.1.1 Interview analysis

To gain insight into the experiences and needs of potential users of a data management platform, part of this paper is dedicated to conducting and interpreting qualitative interviews with five interviewees from different companies and departments, each with different expertise in data and data management. All interviewees were selected based on the buyer personas identified as potential users of the data management platform. The interviews were conducted using a semi-structured script that allowed the interviewer to guide the conversation while allowing for open-ended participant responses.

This interview analysis aims to present the results of the qualitative interviews and provide insight into the perspectives and needs of potential users of the data management platform. Specifically, the analysis will focus on central interview themes, including current data processing, data accuracy and governance, interest in a data governance tool, and data governance challenges.

During the interviews, respondents were asked to imagine a scenario where a document could automatically update its information based on changes made to a parent document. They were then asked to describe how they would envision the feature. Responses were analysed to identify common themes and suggestions. The author will design a wireframe model based on these suggestions in the future chapter.

By exploring these themes, the author aims to identify the key challenges and needs of potential data governance platform users and provide recommendations for designing and developing an effective data management solution.

In the following sections, the author presents a detailed analysis of qualitative interviews using the design thinking methodology described in the theoretical section, the results obtained, and implications for enterprise data management. This analysis provides valuable insights into the experiences and needs of potential users of data management platforms and contributes to the development of effective enterprise data management solutions.

3.1.1.1 Junior Analyst

Category	Description
Positive (+)	The Respondent is focused on entering data into the system and continuously checks it before entering it to ensure it is accurate and up-to-date. Respondent uses matching through patterns or Power Query in Excel to update data, saving time and reducing errors. Respondent suggests an interface/dashboard that shows what data is up-to-date and if it is okay, and a change history feature to see what has changed in the past and who changed it.
Negative (-)	Respondent regularly finds forgotten folders with outdated information, which can lead to errors. Respondent manually updates outdated folders when there is no link to do it automatically, which is time-consuming. Respondent does not use any tool to alert them to changes in different versions of databases or outputs.
Open Questions (?)	The interviewer asked about the process for ensuring data in outputs is still up-to-date when an invoice is updated in the database. The interviewer asked how the Respondent deals with outdated folders when they are found. The interviewer asks if the Respondent thinks it is essential to have detailed information about data lineage and be able to see how the data is linked to other tables. The interviewer asks if there is a platform that the Respondent would like to work with for data governance.
New Ideas (!)	Respondent suggests having rows and columns in updated documents, the data pulled from external databases, and a feature to select whether to update data. Respondent suggests a platform that determines what roles can work with a table and what roles can only work with the table.

Table 9Junior Analyst Interview

Analysis of interviews with respondents shows that they are diligent in their work and take steps to ensure data is accurate and up to date. They use comparison techniques and tools such as Power Query in Excel to update data, save time, and minimize errors. The Respondent also provides insightful ideas to improve data management, such as an interface/dashboard showing current data and a change history feature. However, they need help with forgotten folders and manually updating outdated information, which can be time-consuming and error prone. They also do not use tools to alert them to changes in different versions of databases or issues. The interviewer also asks several open-ended questions about data provenance, legacy components, and the importance of a data management platform.

3.1.1.2 Data Security employee

Category	Description
Positive (+)	The data security employee ensures the data they work with is up to date by correcting data provided by individuals and setting up automatic alerts to check them annually. They track information against public registers and other connected platforms to ensure consistency. They use Microsoft package to keep track of document editing history.
Negative (-)	The data security employee sometimes cannot find needed data in multiple repositories and platforms. They manually search through repositories to find data, which is time-consuming. There is no system in place to notify them of deleted files.
Open Questions (?)	The interviewer asked how the data security employee identifies discrepancies in data and how long it takes. The interviewer asked how they handle outdated information in documents. The interviewer asked if they thought a platform for data governance would be helpful in their work.
New Ideas (!)	The data security employee suggests a platform that allows users to map resources, keep track of who has access to data and documents, and identify who is responsible for the data in a document. They suggest providing training or simpler, user-friendly tools to help users with different levels of technical ability. They also suggest improving data classification and protection, compliance with data protection regulations and standards, and improving how data is shared and transferred between different systems.

Table 10Data Security Employee Interview

The employee ensures data accuracy by correcting data and setting up automatic alerts to check annually. They use Microsoft packages to track document editing history, a good practice for maintaining data integrity. However, the employee sometimes experiences difficulties finding data across multiple repositories and platforms, which can be time-consuming, and there is no notification system for deleted files. The interviewer asked about how discrepancies in data are identified and how outdated information is handled. The employee suggests implementing a platform that can map resources, keep track of data access, and provide user-friendly tools for different levels of technical ability. They also suggest improving data classification, protection, compliance with regulations, and data sharing and transfer between different systems.

3.1.1.3 Decision Maker

Category	Description
Positive (+)	Respondent has a robust system to ensure data accuracy, with accurate tracking of all materials and regular laboratory checks for feed. Respondent also has strict policies in place for data security and uses role-based access control, backup systems, and regular employee education. Respondent suggests a tool that would allow for flexibility in updating reports, with the ability to choose between automatic or manual updates and a version history feature to revert to previous versions of documents.
New Ideas (!)	Respondent suggests a tool notifying users when a parent document changes and allows the option to update the report automatically or manually, depending on the situation. Respondent also suggests having a version history feature to revert to previous versions of documents. To ensure security and privacy, Respondent suggests having different levels of access depending on the user's role within the company and ensuring that any sensitive information is kept confidential and not shared with unauthorized users.
Open Questions (?)	The interviewer asked about the process for ensuring data in outputs is still up to date when an invoice is updated in the database. Respondent: "When an invoice is updated in our database, we have a system that automatically updates all related outputs to reflect the changes. This ensures that our data is always up-to-date and accurate." The interviewer asked how the Respondent deals with outdated folders when they are found. Respondent: "When we come across outdated folders, we have a process to either update them with the latest data or archive them if they are no longer needed. We also try implementing automated systems wherever possible to reduce the chances of outdated folders going unnoticed."
Negative (-)	Respondent feels content with their current systems and processes for managing data and may not be interested in investing in additional IT sources or a data governance platform.

Table 11 Decision Maker Interview

The respondent has a robust system to ensure data accuracy, with regular laboratory checks and strict data security policies. They suggest a tool for flexibility in updating reports and a version history feature. Additionally, they recommend having different levels of access based on user roles to ensure privacy and security. The respondent has an automated system to update outputs when invoices are updated, and they have a process for dealing with outdated folders. The respondent is content with their current systems and may not be interested in investing in additional IT sources or data governance platforms.

3.1.1.4 Hospital employee

Category	Feedback
Positive (+)	The interviewee believes that a data governance platform that connects all hospitals would be outstanding and make their work easier. The interviewee thinks it would be convenient to have a feature that alerts them to incomplete reports that must be uploaded.
Negative (-)	The current platform used by the hospital is ancient and does not offer any text editing features. Editing can be complex, and the lack of mouse support makes it challenging to make changes. The interviewee sometimes needs help finding the data they are looking for when they have yet to finish compiling a report, and someone needs the information urgently.
Open Questions (?)	The interviewer asked if other hospitals have a similar system to the one used by the interviewee, but the interviewee needed clarification. The interviewer asked if the interviewee had ever encountered any company data governance challenges, specifically regarding security and editing permissions. The interviewee mentioned that files can't be edited once closed, but they have experienced difficulties accessing old paper data.
New Ideas (!)	The design thinking exercise involved imagining a scenario where a child document is connected to a parent document so that changes made in the parent document automatically update the child document. The interviewee suggested that the child document should have a link to the parent document, and any changes made in the parent document should be reflected in the child document. They also suggested that if these documents are Word documents, they could be represented as a hyperlink or a coloured box. They also suggested that having old paper data connected to the system would be helpful, even though digitising it would require much work.

Table 12 Hospital Employee Interview

The current data management process used by the hospital has several areas for improvement. The need for a modern platform with text editing features and mouse support makes editing difficult for interviewers. In addition, the interviewer needs help finding the necessary data when they still need to finish composing the report, which can be problematic when someone needs the information urgently. An interesting point that emerged from the interview is the interviewee's use of Microsoft Word to create reports and then copy the text into the old platform used by the hospital. The interviewee explained that editing information in Word is more accessible because the platform does not offer text editing functions. This suggests that the current platform used by the

hospital may need to improve functionality, which could be a significant barrier to effective data management.

Another interesting point is the respondent's mention of difficulties with accessing old paper data. The interviewee states that the hospital has electronic and paper data and that accessing old paper data can be problematic. Although the interviewer believes that connecting the old paper data to the system would be helpful, he also realizes that digitizing all this data would be pretty demanding. This point highlights the importance of considering the technical aspects of a new data management platform and the practical realities of integrating existing data sources and workflows. Feedback from the interview indicates a strong interest in improving the current data management process used by the hospital and that the new platform could bring significant benefits to healthcare providers. However, it also highlights the need to consider practical issues such as data digitization and the importance of incorporating features that address specific problems that users encounter in their daily work. With these factors in mind, a new data management platform could be designed that is both technically efficient and practical, providing tangible benefits to healthcare providers and their patients.

3.1.1.5 Data analyst

Category	Feedback
Positive (+)	The interviewee finds a tool that facilitates data management and governance very helpful. It would make everything easier and faster, especially when manually inputting data. They also think it would help with getting reports to upper management.
Negative (-)	The interviewee needs help with data governance in their company, such as data duplication and getting lost. They sometimes struggle to find the data they need and have to ask colleagues for access. They need help ensuring data accuracy and keeping track of changes made to a document over time.
Open Questions (?)	The interviewer asked how the interviewee envisions integrating a data governance platform into their current work process. The interviewee mentioned that files can't be edited once closed, but they have experienced difficulties accessing old paper data.
New Ideas (!)	During the design thinking exercise, the interviewee suggested that having a platform that connects all departments in the hospital would be outstanding. They also mentioned having a feature that alerts them to incomplete reports that must be uploaded. They also suggested having a tool that automatically updates data information in real-time and a system that tracks and prevents data application. The interviewee said digitizing old paper data and connecting it to the system would be helpful, even though it would require much work.

Table 13 Data analyst Interview

The interviewer describes his current way of working with data, which involves entering and creating data for reports using Excel and SQL databases. They need help ensuring the accuracy and timeliness of the data used, as there are currently no approaches to this. In addition, they sometimes have difficulty finding data and rely on asking colleagues or superiors for help.

Interviewees talk about data management issues in their company, including problems with data loss or duplication, limited access, and difficulties in data retrieval. They indicate that a data management platform that connects all departments would help but incorporating it into their current workflow can be problematic.

The interviewee also describes the potential benefits of a platform that automatically updates information, which could make it easier for them to enter data into reports manually. Overall, the interview highlights the importance of data governance, and the challenges companies can face in ensuring accurate, accessible and up-to-date data.

4 Business case: Implementation of data governance platform into a non-IT department

In the previous chapter, the author presented a business model for developing a data governance platform and mentioned its potential use in enterprises. The qualitative interviews further supported this notion and revealed a growing need for data management solutions across organizations. In particular, one respondent described their ongoing problem with unused data and the need for a data governance platform to address it.

This chapter takes this respondent's case and proposes a business case for developing a data governance platform to solve their data management problems. To ensure confidentiality, the author anonymized the respondent and their organization. Nevertheless, their issues are not unique, and the proposed solution can significantly benefit any organization in a similar situation.

The proposed data governance platform will provide several features to address new data issues, including data lineage mapping, access management based on user roles, data transfer from Excel to the new platform, an intuitive user interface, and the ability to update child documents when parents change files automatically. By developing and implementing this platform, we can help organizations solve data management problems, improve data usage, reduce error rates, and ultimately achieve better business results.

4.1 Executive summary

The non-IT department currently needs help managing data effectively. Department staff are not IT experts and store data in multiple Excel files in various folders, resulting in duplicate or outdated information and potentially incorrect conclusions. To address these issues, we propose implementing a data governance platform to map data provenance, manage access based on user roles, and provide a user-friendly interface to retrieve information quickly. The proposed data governance platform is specifically selected to meet the needs of non-IT staff with a user-friendly interface and intuitive features that enable them to work with data more efficiently. By migrating data from Excel to the new platform, we can ensure that everyone in the department works with the same accurate information regardless of their IT expertise.

A decisive feature of the proposed platform is the ability to update child documents when parent files change automatically. This feature is especially important in non-IT departments, where staff may need more time or expertise to update each file when changes are made manually. Automating this process can save time and reduce errors, leading to better decision-making and business outcomes.

A data governance platform can also help departments save money in several ways. First, by making it easier to find and use data, employees can spend less time searching for information and more time on productive tasks. This can increase efficiency and productivity and, ultimately, lower costs. In addition, by reducing the risk of errors and misunderstandings, the platform can help prevent costly mistakes and improve the accuracy of reports and other deliverables.

For example, a department needs to create a report for management that relies on data from multiple Excel files. With a data management platform, staff can manually review each file to ensure the data is accurate and current. This process can take several hours or even days, depending on the complexity of the report. With the proposed platform, staff can quickly and easily access the data they need and have confidence that it is accurate and up-to-date. This can save significant time and reduce the risk of errors, ultimately improving report quality and overall department performance.

4.2 Product / Service description

The proposed data governance platform provides an innovative and comprehensive solution for organizations that need to manage and leverage large amounts of data. The platform incorporates the best features of existing data management platforms. It adds a new innovative element: a pipeline-based approach automatically updating child documents or databases based on parent document or database changes. This feature streamlines workflows and increases the accuracy and quality of data management, setting it apart from the competition.

The platform includes various data management features such as collaboration tools, pipeline management, quality management, and cataloguing. The platform's data cataloguing capabilities enable organizations to collect and catalogue data assets from various sources and provide

comprehensive metadata and contextual information about each asset. Natural language searches and browsing capabilities simplify the searching and discovery of data assets, making it easier for users to find and use data.

Data line capabilities provide users with information about the end-to-end data line of the platform, enabling organizations to make better data-driven decisions, manage data risk, and ensure regulatory compliance. The platform's data quality management capabilities allow users to create rules and policies and automate data validation and auditing to ensure data reliability and accuracy.

User interface is simple and intuitive, making it easy for users to find and use data sources. The platform offers a selection of dashboards and reports that provide insight into data sources, usage, and quality, making it easier for users to make informed decisions about data usage.

The unique feature of the proposed data management platform, which includes automatic updating of child documents or databases based on changes in the parent document or database, is precious for organizations that need to manage large volumes of data. For example, this feature can be used in a manufacturing company to automatically update all related child documents when changes are made to the original drawing or design of an automotive part to ensure that all data is accurate and up-to-date, reducing costs, risk of error, and improving the manufacturing process.

4.3 Project Organization

A clear project organization is essential to see a data management platform's comprehensive development and implementation. The project organization encompasses several roles and responsibilities, from the project sponsor, who provides funding and resources, to the technical development team that develops the platform, to the change management team that manages the organizational changes associated with the project. A clear project organization allows the project team to work together effectively to achieve the project goals.

Role	**Responsibilities**
Project sponsor	Provides funding and resources, ensures the project aligns with department goals and communicates with stakeholders and executives.
Project manager	Oversees project plan, works with the team to ensure the project stays on track and meets objectives, and reports progress to the project sponsor.
Business analyst	Works with stakeholders to define data governance policies, develop use cases, identify requirements, and translate business requirements to technical specifications.
Technical lead	Oversees platform architecture selects appropriate technologies and tools, and manages technical development team.
Data governance manager	Develops and implements data governance policies and procedures, defines data standards, establishes data quality rules, and creates data catalogues.
Data Architect	Designs data architecture for the platform ensures that data is modelled effectively and efficiently and establishes data integration and flow standards.
Technical development team	Develops and implements data governance platform, using specs provided by the business analyst and technical lead.
Testers	The test data governance platform develops test plans, performs functional and performance testing, and provides feedback to the technical development team.

Table 14 Business case project roles

4.4 Financial Appraisal

The objective of this financial assessment of a data governance platform project is to provide a quantitative analysis of the costs and benefits associated with the development and implementation of the platform. The main prices considered in the assessment include the salaries of the team members who will participate in the project and the licence of the data management platform. This assessment determines whether the project is financially feasible and can deliver a positive return on investment (ROI). By conducting a thorough financial analysis, we can identify the costs of platform development and the expected benefits, including potential cost savings, improvements in data quality and efficiencies that will support the decision-making process and ensure that the project is aligned with the strategic objectives of the non-IT department.

4.4.1 Project Plan

Before starting a financial appraisal, a project plan is vital because it provides detailed information about the timeline and other essential aspects of the project. This information is necessary for an accurate financial appraisal, which involves analyzing the project's costs, benefits, and risks. With a project plan, estimating costs and benefits and identifying potential risks and challenges is easier. A well-developed project plan can also help stakeholders and investors understand the task better and make informed decisions about funding and resource allocation.

The WBS table summarises the project plan into key phases and associated tasks. The table is organized by stage and includes a list of the specific functions that will be completed as part of each phase.

Phase	Task
Project Initiation Phase	Define project scope, objectives, and success criteria. They form a project team and allocate roles and responsibilities, Conduct stakeholder analysis and engagement planning, Define project timeline and budget
Requirements Gathering Phase	Identify and document requirements for the data governance platform. Conduct user research to understand user needs and workflows. Define data governance policies and procedures
Vendor Selection Phase	Conduct market research to identify potential vendors Develop request for proposal (RFP) Evaluate vendor proposals and select preferred vendor Negotiate a contract with the selected vendor
Implementation Phase	Install and configure data governance platform Develop and execute data migration plan Develop and execute user training plan Conduct user acceptance testing (UAT) and resolve any issues Roll out the platform to non-IT department
Post-Implementation Phase	Conduct a post-implementation review to assess project success. Identify and resolve any issues. Develop and implement ongoing maintenance and support plan Plan for future upgrades and enhancements

Table 15 Business Case Phases

The table below shows the activities' Early Start (ES), Early Finish (EF), Late Start (LS), Late Finish (LF), and Slack.

Activity	LS	ES	LF	EF	Slack
1	7	0	7	7	0
1.1	2	0	2	2	0
1.2	4	0	4	4	0
1.3	2	0	2	2	0
1.4	6	4	8	8	0
2.1	9	6	11	9	2
2.2	8	4	8	8	0
2.3	8	5	9	8	1
2.4	7	3	7	7	0
3.1	10	3	14	9	5
3.2	18	7	18	14	4
3.3	13	5	13	10	3
3.4	16	6	16	12	4
4.1	26	8	26	19	7
4.2	11	6	13	11	2
4.3	11	6	11	10	1
4.4	17	9	17	13	4
5.1	20	13	20	16	4
5.2	32	16	32	28	4
5.3	24	14	24	20	4
5.4	20	11	20	16	4
6					NA

Table 16 Business Case Critical Path

Based on the calculations in the tables in the appendix and the definition of the critical path, the critical path for this project is 1-1.4-2.2-2.3-3.2-4.1-5.2-6.

The total duration of the critical path activities is:

4.33 + 4 + 4 + 4 + 7 + 8 + 6 = 32 weeks

Since the critical path activities have no slack, any delay in these activities will directly delay the completion of the project. Therefore, the total project duration is 35.33 units of time, the same as the critical path duration.

Week	Activity Number	Activity Description
1	1.1	Define project scope, objectives, and success criteria
2	1.2	Form a project team and allocate roles and responsibilities
	1.3	Conduct stakeholder analysis and engagement planning
3	1.4	Define project timeline and budget
4-7	2.1	Identify and document requirements for the platform
	2.2	Use research to understand user needs and workflows
8-9	2.3	Define data governance policies and procedures
10-11	3.1	Conduct market research to identify potential vendors
12-13	3.2	Develop a request for proposal (RFP)
14-15	3.3	Evaluate vendor proposals and select the preferred vendor
16	3.4	Negotiate a contract with the selected vendor
17-22	4.1	Install and configure the data governance platform
23-25	4.2	Develop and execute a data migration plan
26-27	4.3	Develop and execute a user training plan
28-29	4.4	Conduct user acceptance testing and resolve any issues
30	5.1	Conduct post-implementation review
31	5.2	Identify and resolve any issues
32-33	5.3	Develop and implement maintenance
34-35	5.4	Plan for future upgrades and enhancements
36	6	Project closeout and documentation

Table 17 Business case activity in weeks

The table represents a project schedule that outlines the activities required to complete a data governance project. The schedule is divided into 36 weeks, with each activity assigned a specific week in which it is expected to be completed. The activities are listed in order based on their logical sequence, with some activities being dependent on the completion of others.

4.4.2 Identifying costs

Position Name	ISPV	Wage	MD Cost	Number of employees	Sum of wages	Allocated MDs	The sum of costs per role
Project Manager	12232 Technical development managers	107,125 Kč	4,869 Kč	1	107,125 Kč	100	486,930 Kč
Business Analyst	26312 Business economics specialists	71,405 Kč	3,246 Kč	1	71,405 Kč	100	324,567 Kč
IT Manager	13302 Information technology and operations managers	130,820 Kč	5,946 Kč	1	130,820 Kč	100	594,634 Kč
Database Administrator	2521 Database designers and administrators	75,726 Kč	3,442 Kč	2	151,451 Kč	50	344,207 Kč
Technical Writer	33391 Marketing, promotion and advertising personnel	39,155 Kč	1,780 Kč	1	39,155 Kč	5	8,899 Kč
Customer support specialist	33391 Marketing, promotion and advertising personnel	39,155 Kč	1,780 Kč	1	39,155 Kč	5	8,899 Kč
System Administrators	2522 System administrators, computer network administrators	69,790 Kč	3,172 Kč	2	139,580 Kč	50	317,227 Kč
UI Designer	2521 Database designers and administrators	75,726 Kč	3,442 Kč	2	151,451 Kč	50	344,207 Kč
QA Engineers	2519 Software testing and related professionals	81,725 Kč	3,715 Kč	2	163,450 Kč	15	111,443 Kč
HR Manager	2424 Training and human resources development specialists	63,359 Kč	2,880 Kč	1	63,359 Kč	50	143,997 Kč

Table 18 Business Case costs

The Business Case costs table provides information on the individual roles involved in the project, their average ISPV salaries for the first half of 2022, and the number of employees required for each part.The list of job titles provided has been translated from Czech to English, as it originally appeared in the 2022 first-half report in ISPV.

It also includes costs allocated to business leaders, total wages and total costs per role. The spreadsheet is helpful for financial evaluation because it provides a better understanding of the

labour costs associated with the project. Developing an accurate budget and ensuring the project's financial feasibility is necessary.

It should be noted that when calculating individual role costs, it is assumed that employees assigned to a project are not assigned to that project on a full-time equivalent (FTE) basis. Instead, it is assumed that these employees are the organization's full-time employees and work on other projects or tasks. The assigned MD (person days) column considers this and represents calculations for the full FTE of time each employee is expected to spend on the data management implementation project. With this table, we expect **2,685,011 Kč** to be allocated to salary costs for this project.

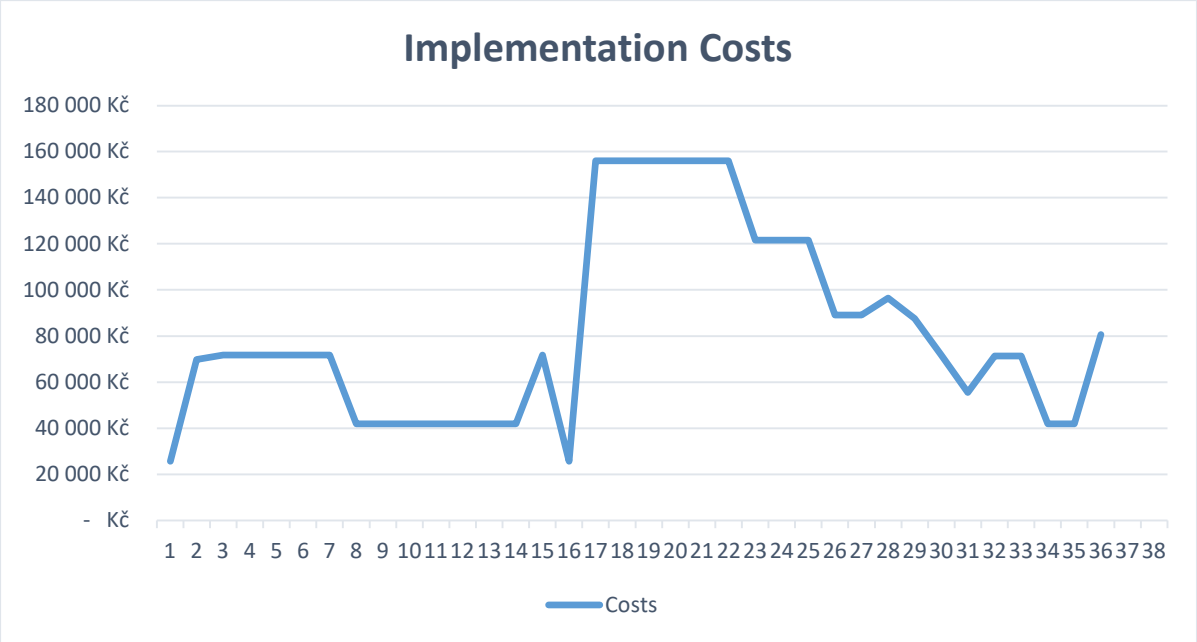


Figure 5Business Case Implementation Costs

The line chart shows the cost evolution of implementing the data governance platform project over 36 weeks. The chart shows that costs increase non-gradually from week 1 to week 23, with a peak in spending during weeks 20 to 23. After this peak, prices decrease and reach their lowest point during the first and last phases of the project. This pattern in cost trends is due to the different phases of the project. During the initial phase, costs are relatively low because most planning and scoping activities are being done, and the team size is smaller. As the project moves into the implementation phase, costs increase as more resources are required for system configuration, data

migration, and user training. The peak cost will likely be in weeks 20 to 23, as this is when most of the implementation work occurs. As the project moves into the post-implementation and final phases, the team size and resources required gradually decrease, resulting in a decrease in cost.

It is important to note that the costs include employee and platform license costs, with license costs of 1098.57 Kč (conversion from 49 USD) per user when the number of users must be at least 5. The line chart visually represents project costs over time, which can help monitor the budget and adjust as needed.

A detailed description of project costs can be found in the table in the attachments of this work.
The total costs after adding the licence costs are estimated to be **2,942,319 Kč**.

4.4.3 Two years cost plan

The chart shows a plan for the costs of a data governance platform over two years, broken down by month. The first eight months include the implementation costs and licenses, while months 9 to 24 include only licenses and customer support with an expected decreasing workload and decreasing costs.

The cost per license per user decreases from the first month when there are more than 100 users to 650 Kč per user. In the first month, there are only five users, and the cost per user is 1,098.57 Kč. This results in a total cost of 317,201.70 Kč, which includes implementation costs.
In the following months, there are still five users, and the cost per user remains the same, resulting in total costs of 275,795.32 Kč, 246,063.60 Kč, and 281,253.01 Kč for months 2, 3, and 6, respectively.

There are seven users in months 4, 5, and 7, resulting in a higher total cost of 658,570.54 Kč, 685,604.24 Kč, and 363,889.41 Kč, respectively.
Months 8 to 24 see a gradual increase in users, resulting in higher total costs. However, the cost per user decreases when there are more than 100 users, resulting in a lower total cost. For example, in month 9, there are 20 users, and the cost per user is 1,098.57 Kč, resulting in a total cost of 42,647.71 Kč. In month 13, there are 100 users, and the cost per user decreases to 650 Kč, resulting in a total cost of 78,701.82 Kč.

A detailed description of project costs can be found in the table in the attachments of this work. The total costs after adding the licence costs are estimated to be **4,519,704.70 Kč.**

Figure 6 Business Case Total Costs

4.4.4 Estimated savings

It is easier to estimate the savings that can be achieved by implementing a data governance platform by providing more information about the specific processes and issues that the department faces and that the company considers sensitive information that cannot be provided for this work. However, here are some potential areas where implementing a data governance platform could lead to cost savings.

- **Better tracking and reporting of expenses** could lead to more efficient resource management and lower overall costs.
- **Optimized property and travel management processes** could reduce administrative costs and increase efficiency.
- **Better decision-making based on quality data** could lead to cost savings through more efficient resource allocation and risk management.
- **More effective contract and supplier relationship management** could lead to cost savings through more favourable terms and better negotiations.

It should be noted that cost savings depend on several factors, including the current state of data management and the specific challenges facing departments. Therefore, a detailed analysis of

departmental operations would be required to determine the typical cost savings achieved by implementing a data management platform. Here are some very rough estimates of how much money in Kč could be saved annually by implementing a data management platform:

- **Improving data accuracy:** the company could save an estimated 0.5 million Kč annually by reducing property and travel data errors.
- **Automation of manual processes:** If data processing costs were reduced by automating manual processes related to the management of property and roads.
- **Increased efficiency:** The bank saved an estimated 1 million Kč annually by streamlining work procedures and eliminating redundant tasks.
- **Better contract management:** Using a data governance platform to manage contracts with suppliers, the company could negotiate better terms and avoid penalties, saving an estimated 0,75 million Kč annually.
- **Better compliance:** The company could avoid fines and penalties by ensuring compliance with legal and regulatory requirements, saving an estimated 0.1 million Kč annually.

These are rough estimates based on the limited information provided, and savings may vary depending on specific conditions.

4.4.5 Return on Investment

First, it is important to note that the Return on Investment (ROI) presented here is a rough estimate, as we cannot publish sensitive information from the department that would allow for a more precise calculation. However, based on the information available, we can see a positive ROI for the data governance project over several years.

Looking at the chart below, we see that in the first year, there were costs of 3,295,338 Kč, and no savings or returns were realized. This is expected, as the initial implementation and setup of the data governance platform incur some costs, but the benefits are not yet fully realized.

However, in the following years, we see a significant number shift. By the second year, there were savings of 1,175,000 Kč, which offset some of the costs incurred in the first year, resulting in a

negative return of -3,344,705 Kč. This trend continues into the third and fourth years, with increasing savings and decreasing costs resulting in smaller negative returns.

By the fifth year, we see a complete turnaround, with savings of 2,350,000 Kč over the costs of 1,224,367 Kč, resulting in a positive return of 32,194 Kč. This trend continued into the following years, with increasing savings and positive returns realized.

It is important to note that the positive ROI reflects the cost savings achieved by implementing the data governance platform. These cost savings come in the form of increased efficiency, reduced errors, and better management of resources.

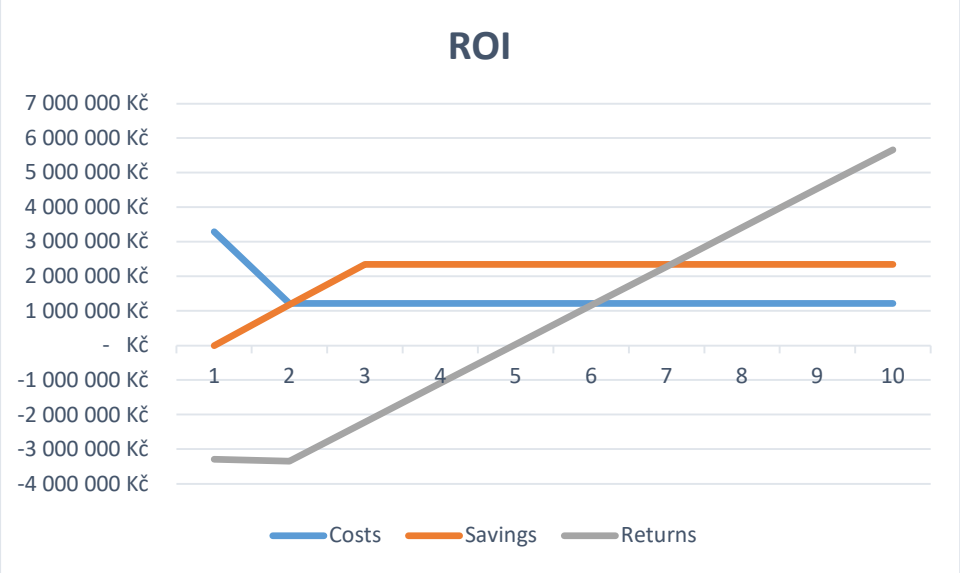


Figure 7 Business Case ROI

4.5 Risk Assessment

Implementing a new data governance platform involves several risks that must be carefully considered and managed for a successful outcome. These risks can generally be divided into five main areas: technical, operational, financial, data security, and legal and regulatory.

One significant technical risk is the possibility of compatibility issues between the new platform and existing software. This could result in delays or errors in implementation and additional costs

to fix any problems. In addition, the platform may require significant technical expertise to operate, which could lead to delays and errors.

Another significant risk is operational. Employees using the new platform may need help with a learning curve, which could lead to errors or delays in departmental workflow. In addition, the implementation process may require a significant amount of time and resources from departmental staff, which may impact their ability to perform their regular duties.

The financial risk associated with implementing a new platform can also be significant. The cost of the platform and its implementation may be higher than anticipated, and additional ongoing charges for training, maintenance and updates may be incurred. This could impact the department's budget and financial performance.

Another critical area of risk to consider is data security. The platform may threaten departmental data's security and privacy, mainly if it integrates with other systems or involves third-party data storage. This could lead to data breaches, damage to the department's reputation and high financial and legal costs.

Last but not least, legal and regulatory risks must be carefully managed. Implementing a data governance platform may require compliance with various legal and regulatory requirements, such as data protection laws. Failure to comply with these requirements could result in financial penalties or legal proceedings, damaging the department's reputation and financial performance.

To effectively manage these risks, it is necessary to carefully evaluate them and develop strategies for mitigating or gathering them as part of the implementation process. This may include careful planning, staff training and support, regular platform reviews and testing, and ensuring compliance with the platform and its use with legal and regulatory requirements.

Risk	Likelihood (1-10)	Impact (1-10)	Mitigation
Data breaches	8	9	Conduct thorough data security assessments, implement access controls, encrypt sensitive data, and train employees on data security best practices.
Non-compliance with regulations	7	8	Stay up-to-date on relevant laws and regulations, implement policies and procedures to ensure compliance, and conduct regular audits to ensure adherence to regulations.
Data quality issues	5	6	Develop and implement data quality standards, establish clear ownership and responsibilities, and conduct regular data quality checks.
Resistance to change	4	7	Communicate the benefits of the data governance platform, involve employees in the implementation process, provide sufficient training and support, and address any concerns or feedback.
Technical difficulties	6	8	Conduct thorough testing before implementation, have a contingency plan in case of technical issues and have a dedicated technical support team.

Figure 8 Business Case Risk

4.6 Business case conclusion

After analyzing the business case for implementing a data governance platform and reviewing the proposal for a specific platform designed for non-IT staff, it is clear that such a platform can significantly positively impact departmental performance. The proposed platform provides staff with a user-friendly interface, allowing them to manage and query data efficiently, saving time and increasing productivity. Migrating data from Excel to the new platform ensures that everyone works with the same accurate information, reducing the risk of errors and misunderstandings. Additionally, the platform's ability to automatically update child documents when parent files change further reduces errors and saves time. However, the potential risks associated with implementing a data governance platform, such as data breaches, compliance issues, and employee resistance, careful planning, training, and a well-defined risk management strategy can mitigate these risks. The benefits of implementing a data governance platform far outweigh the risks. In addition to improving departmental performance, the platform can save the company money by increasing efficiency and productivity and reducing the risk of errors and misunderstandings. The platform also sets a precedent for the rest of the organization, improving data management across departments and potentially leading to improved business outcomes. Based on the information provided throughout this thesis, implementing a data governance platform in the non-IT department of a multinational bank could lead to significant cost savings and process improvements. The initial costs for implementation and licenses are expected to be high but can be recouped within the first year, followed by consistent savings in subsequent years.

5 Use cases

Based on the interviews in chapter 3.1.1 following use cases were proposed.

5.1.1.1 Junior Data Analyst adding contextual information to data/document.

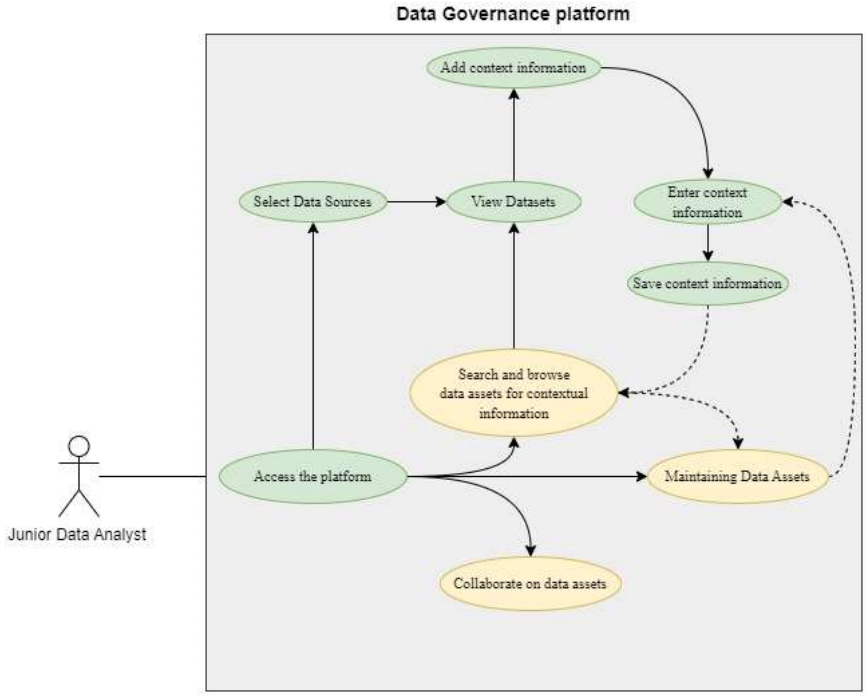


Figure 9 Junior Data Analyst adding contextual information to data/document

Access the platform: Junior data analysts should open their web browser and navigate to the data governance platform login page. They should click the login button and enter their credentials to access the platform.

Select Data Sources: Junior Data Analyst should click on the Data Sources tab from the main navigation menu at the top of the page. A drop-down menu with a list of data sources they can access will appear. They should select the data sources that contain the datasets they want to add contextual information to by clicking the checkboxes next to their names.

View Datasets: After the Junior Data Analyst has selected the data sources, they should click the Datasets button on the main navigation menu. This will take them to the page showing

all the data assets in the selected sources. They should click on an asset's name to view its details.

Add context information: Junior Data Analyst should scroll down to the Context Information section of the data object detail page. They should click the "Add" button on the right of the "Context Information" heading to add new context information.

Enter context information: In the "Add context information" pop-up window that appears, the junior data analyst should enter the relevant information about the data object, e.g. B. Purpose, relevance and importance to the organization. They can also add tags or labels to classify the data object by category or keyword. After entering the information, they should click the "Save" button at the bottom of the pop-up window to save the context information in the database.

Save context information: The platform refreshes the page, and the junior data analyst sees the added context information in the context information section of the data object detail page. If They need to add more context information, repeat steps 4 through 5 for each information They want.

View context information: The junior data analyst should scroll down to the context information section on the data object detail page to view the added context information. They will see a list of all the contextual information they have added, the date, and the user who added it. To edit or delete the information, they should click on the "Edit" or "Delete" buttons next to each piece of information.

Search and browse data assets for contextual information: Junior Data Analyst should click the Search and Browse tab in the main navigation menu to search and search data assets based on their metadata and context information. This takes them to the search and browse page, where they can use the platform's natural language search and browsing capabilities to locate data assets. They can also use the filters or facets to narrow the search results based on specific criteria, including the added context information.

Collaborate on data assets: To collaborate on data assets with other users, the Junior Data Analyst should click on the Collaboration tab in the main navigation menu. This takes them to the collaboration page, where they can share data assets with other users, collaborate on data-related tasks, and provide feedback or comments on data assets.

Maintaining Data Assets: To maintain data assets and ensure their accuracy, relevance and importance to the business, the Junior Data Analyst should click on the Maintenance tab in the main navigation menu. This takes them to the maintenance page, where they can update the metadata and contextual information of assets, validate their quality, and adjust access controls.

5.1.1.2 Data Security employee validate the information in the dataset.

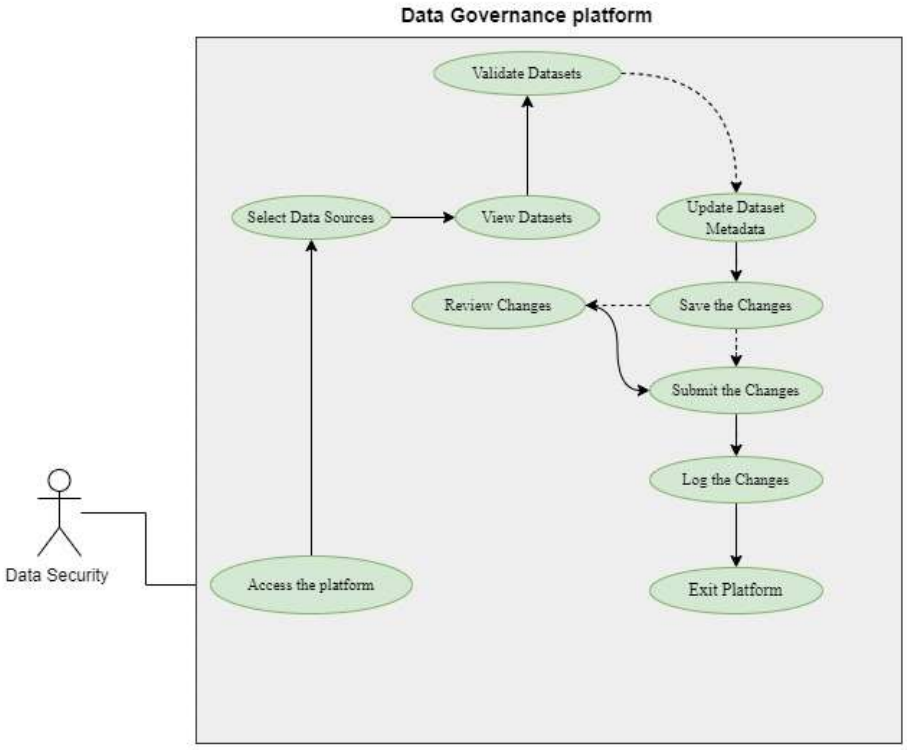


Figure 10 Data Security employee validate the information in the dataset

Access the Platform: The data security employee should open their web browser and navigate to the data governance platform login page. They should click the login button and enter their credentials to access the platform.

Select Dataset: The data security employee should select the dataset that needs to be validated. They should click the Datasets button on the main navigation menu. This will take them to the page showing all the data assets. They should click on the asset's name to view its details.

View Dataset Details: After selecting the dataset, the data security employee should scroll down to the Dataset Details section of the page. Here, they can view the dataset's metadata and other details.

Validate the Dataset: The data security employee should check the metadata and other details to validate the dataset. They should click the Edit button next to the metadata to modify it if necessary.

Update Dataset Metadata: In the pop-up window that appears, the data security employee should update the metadata, such as the dataset name, description, and other relevant information. They should click the Save button to save the changes.

Save the Changes: The platform refreshes the page, and the data security employee can see the updated metadata in the Dataset Details section.

Review Changes: The data security employee should review the changes made to the dataset to ensure that the metadata is accurate and up-to-date.

Submit Changes: If the data security employee is satisfied with the changes, they should submit them by clicking the Save button at the bottom of the page.

Log Changes: The platform logs the changes made to the dataset, the date and time of the change and the user who made it. This ensures a record of all changes made to the dataset.

Exit Platform: Once the data security employee has completed their validation task, they should log out of the platform to ensure the security of the data.

5.1.1.3 Hospital employee adding patient information to a medical record

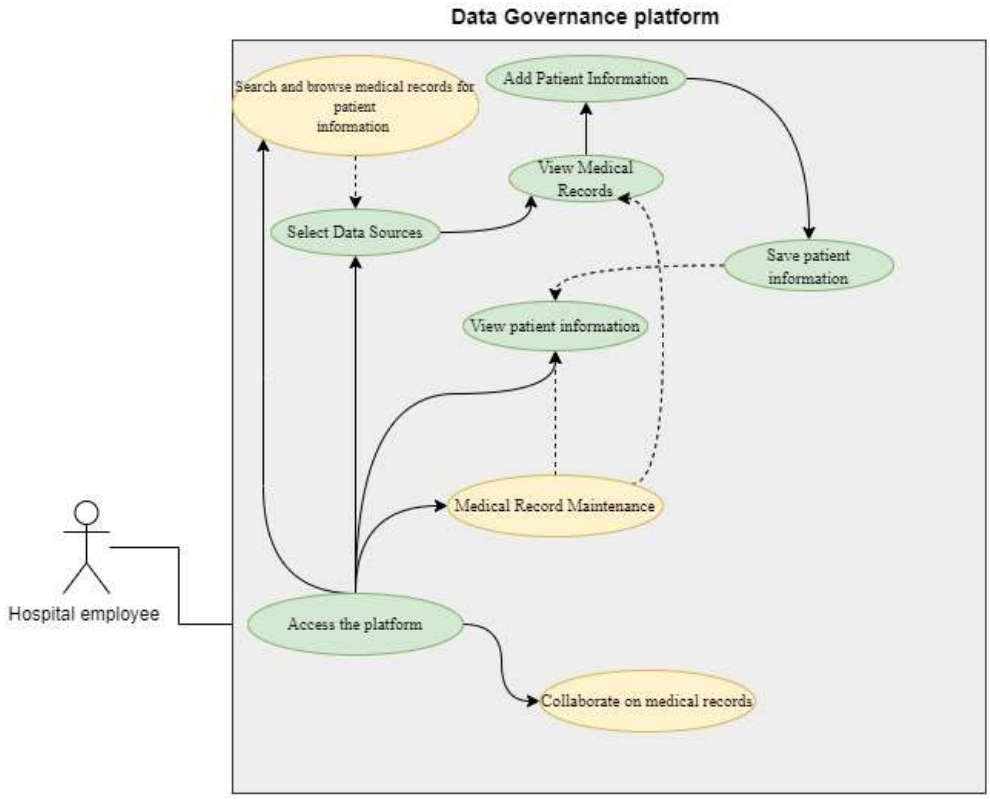


Figure 11Hospital employee adding patient information to a medical record

Access to the platform: The hospital employee should open their web browser and go to the login page of the data management platform. He should click the login button and enter his credentials to access the platform.

Select data sources: A hospital employee should click on the Medical Records tab in the main navigation menu at the top of the page. A drop-down menu with a list of health record sources they can access will appear. They should select the resource containing the patient's medical record to which they want to add information by clicking the check box next to its name.

View Medical Records: After a hospital employee selects a medical records source, they should click the Medical Records button in the main navigation menu. This will take them to a page showing all the medical records in the selected resource. They should click on the name of the patient whose medical records they want to update and view their details.

Add patient information: The hospital employee should go to the Medical Records Details section on the medical record details page. He should click the "Add" button on the right side of the "Patient Information" header to add new patient information. Enter patient information: In the "Add patient information" pop-up window, the hospital employee should enter relevant patient information, such as B.'s vital signs, test results or medication history. It can also add tags or labels to organize data by category or keyword. After entering the information, he should click the "Save" button at the bottom of the pop-up window to save the information to the database.

Save patient information: The platform refreshes the page, and the hospital employee will see the added patient information in the Medical Record Details section of the Medical Record Details page. If he needs to add more patient information, he should repeat steps 4 to 5 for each piece of information required.

View patient Information: The hospital employee should scroll down to the Medical Record Detail page's Medical Record Details section to view the added patient information. He will see a list of all added patient information, the date and the user who added it. If he wants to edit or delete the data, he should click the "Edit" or "Delete" buttons next to each piece of information.

Search and browse medical records for patient information: Hospital employee should click the Search and Browse tab in the main navigation menu to search and browse medical records based on their metadata and patient information. This will take him to a search and browse page where he can use the platform's natural language search and search capabilities to find medical records. They can also use filters or facets to narrow search results based on specific criteria, including added patient information.

Collaborate on medical records: Hospital employee should click the Collaboration tab in the main navigation menu to collaborate on medical documents with other users. This will take them to a collaboration page where they can share health records with other users, collaborate on health-related tasks, and provide feedback or comments on health records.

Medical Record Maintenance: To maintain medical records and ensure they are accurate, relevant and vital to the business, the hospital employee should click the Maintenance tab in the main navigation menu. This takes him to a maintenance page where he can update health record metadata and patient information, verify their quality, and modify access controls

5.1.2 Use cases Summary

Role	Purpose	Key Tasks	Specialization
Junior Data Analyst	Adding contextual information to data/documents	1. Navigate data sources 2. Select datasets 3. Add and save context information 4. Search and browse data assets 5. Collaborate and maintain data assets	Enriching data with context for better understanding and decision-making
Data Security Employee	Validating information in datasets	1. Access platform 2. Select dataset 3. Check and update metadata 4. Review changes 5. Log changes	Ensuring accuracy and relevance of dataset information for data quality and security
Hospital Employee	Adding patient information to medical records	1. Navigate data sources 2. View medical records 3. Add and save patient information 4. Search and browse medical records 5. Collaborate and maintain medical records	Maintaining accurate and up-to-date patient information in healthcare settings

Table 19 Use Cases summary

6 Wireframe

The wireframe described in this document was proposed based on design ideas generated from qualitative interviews conducted with potential users of the data governance platform. These interviews were analysed in the previous chapter to identify key themes and pain points experienced by users in managing their data. The wireframe was designed to address these pain points and incorporate the desired features and functionalities requested by the respondents. This approach ensured that the wireframe was user-centred and optimized for the needs of its intended users.

The wireframe was developed with a deep understanding of user needs and preferences by employing the design thinking method in the interview process, resulting in a more effective and user-friendly data governance platform.

The depicted wireframe serves to visualize this feature's framework containing several elements essential for effective data management.

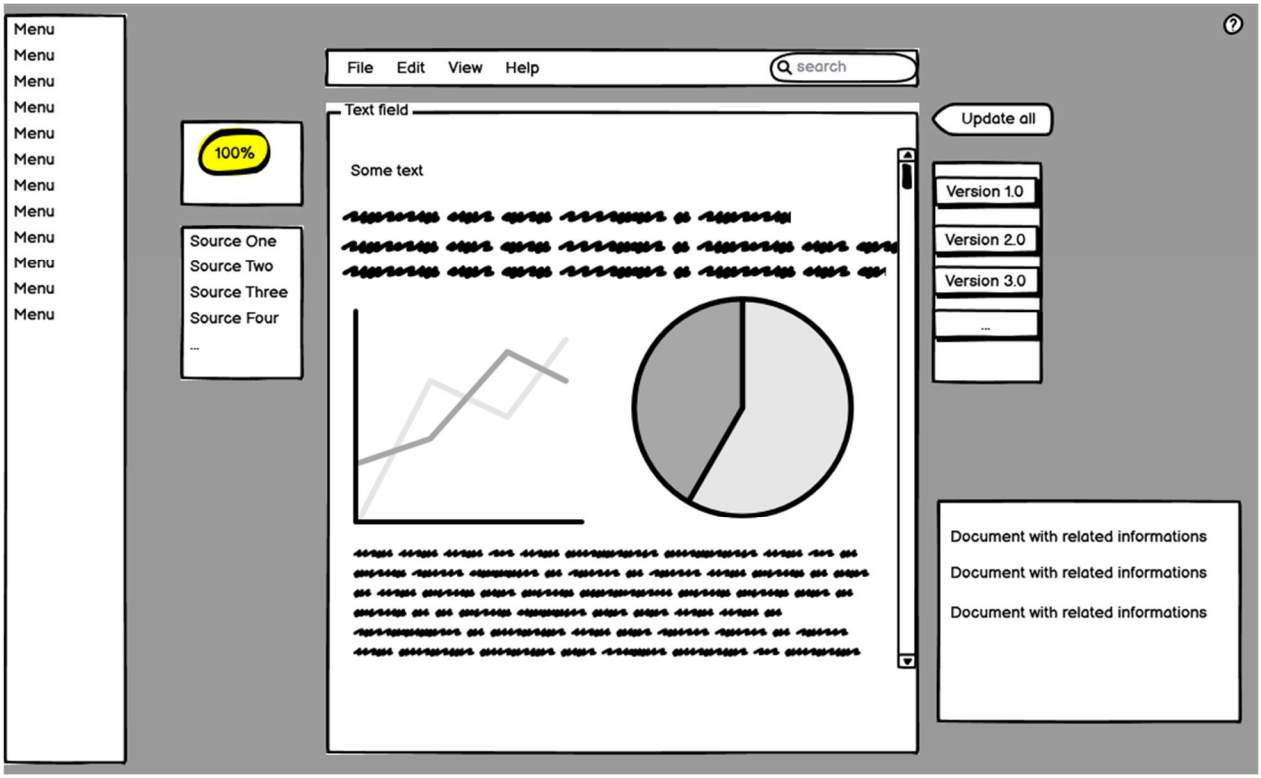


Figure 12 Wireframe screen view

At the top of the screen is a horizontal menu with several bars related to document editing. These bars provide users with options for creating new documents, opening existing documents, cutting/copying/pasting text, formatting text, and inserting images or other media into documents.

In addition to document editing and platform-related bars in the horizontal menu, users can link data to other documents, tag data, and more. These additional features allow for greater flexibility in managing and organizing information within the platform.

On the right side of the screen is a vertical menu that contains bars related to the platform itself. These bars allow users to navigate different platform parts, view their documents, adjust settings, and access helpful resources when needed. This menu is the same for all pages and features of the platform.

At the top left corner of the screen is a percentage field that updates in real-time as users make changes to their documents. If the percentage falls below 100%, some information in the document needs to be updated. This feature ensures that users always have access to the most up-to-date information.

The feature of having a real-time percentage field at the top left corner of the screen was developed after analysing an interview with a junior data analyst (chapter *3.1.1.1 Junior Analyst*). The interviewee suggested that an interface or dashboard that shows the status of the data, whether it is up-to-date or not, would be helpful. Additionally, they recommended a change history feature to track what changes have been made to the data in the past and who made them. Based on these insights, the real-time percentage field was incorporated into the wireframe to ensure that users always have access to the most up-to-date information. If the percentage falls below 100%, the users are alerted that some knowledge in the document needs to be updated. This feature is expected to enhance the overall user experience by providing a simple and effective way to monitor the status of their papers.

Below the percentages field is the source files field, which lists all the source files used to generate the information in the document. Users can click on any of these files to view additional information about the resource, such as its name, location, and last update time. This feature is critical to informing users of where their information comes from. This feature was implemented after an analysing interview with a Data security employee *(chapter 3.1.1.2 Data Security employee)*

There is an "update all" button in the upper right corner of the screen, which updates all the information in the document at once by loading the most recent information from the source files. This can be a valuable feature for users who want to ensure that their document is fully up-to-date quickly.

Below the "update all" button is version buttons allowing users to view their document's previous versions. When the user clicks the version button, the text in the document is updated to show

information from that version. This can be useful for comparing changes between versions or referencing information from older document versions.

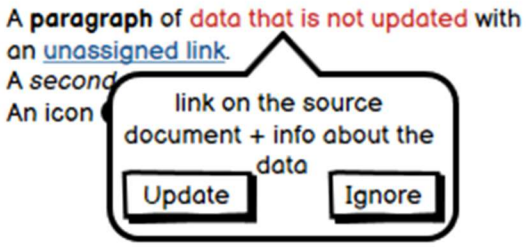


Figure 13 Wireframe text view

When the user moves the cursor over the highlighted text in the document, a window will appear with additional information about the source of information. This may include details such as the source file's name, the type of data used and the time of the last update. The "Update" button in the hover window allows users to update the information for that data. In contrast, the "Ignore" button allows them to ignore it and keep the existing data in the document. This feature is critical for users to make informed decisions about the data they use.

Within the text, specific values taken from another document are highlighted. These values are colour-coded to provide users with visual cues about their freshness. Green indicates the information is current, while orange indicates it uses older information from its parent document. Red signifies that the information is being pulled from a parent document with outdated information. This feature is critical for helping users quickly identify and update any outdated information within their documents.

7 Results and Discussion

7.1 Conclusion and suggestions for data governance platform

The proposed data governance platform underwent several strategic analyses using a business model canvas, SWOT, TOWS, and Porter's Five Forces Analysis.

The analysis identified the strengths of the proposed data management platform, including a team of skilled and experienced developers and IT professionals, collaboration with academic experts, and a unique feature that automatically updates data in "child documents." The platform also faces several challenges, such as the need for brand recognition, reliance on external funding, and external threats, such as cybersecurity risks and the rapid pace of technological change.

One interesting finding from the analysis is the high threat of new entrants in the data management market due to low barriers to entry and growing demand for data management solutions. Established players with significant market share and brand recognition may deter some new entrants. However, smaller players can challenge incumbents through innovation and differentiation.

Another notable finding from the analysis is the importance of collaborating with academic institutions to develop innovative solutions and stay ahead of emerging trends. The innovative approach to automatically updating child documents of the proposed platform is an example of such collaboration that can streamline workflows and improve data accuracy.

Given the rapid pace of technological change and cybersecurity risks, the analysis also highlighted the need for agility and adaptability in the data management market. Hack SolarWinds highlights the significant risks and vulnerabilities associated with data management platforms. The proposed platform must be agile and adaptable to new trends and threats to remain relevant and competitive.

The TOWS matrix developed from the SWOT analysis highlights the need for the platform to focus on building brand awareness, developing strategic partnerships and integrating new technologies into its offerings to overcome the identified weaknesses and threats.

The market analysis highlights the importance of data accuracy and quality, a significant issue for many businesses in the Czech Republic and beyond. An innovative approach to the data feed of the proposed platform could provide a significant advantage in addressing this issue and increase the platform's attractiveness to potential customers.

The strategic analysis of the proposed data governance platform provided a comprehensive understanding of its business model, internal and external environment and competitive landscape. The analysis results can be used to develop effective strategies to overcome potential challenges and exploit opportunities in the market. Identified benefits highlighted the importance of collaborating with academic experts, building brand awareness, integrating new technologies, and investing in cybersecurity measures to remain competitive in the rapidly evolving data management market.

7.1.1 Data governance platform development schedule and time analysis

The proposed project involves the development of a data governance platform that consists of several activities that need to be completed in a specific sequence. These activities include project planning, requirements gathering, design, development, testing and deployment. Each phase consists of several sub-activities, and the duration of each activity is estimated based on expert judgement, historical data and other relevant factors.

The analysis provides a detailed breakdown of the activities and paths involved in the project and their expected durations and variations. By identifying the critical path, project managers can focus on ensuring that activities are completed on time and with minimal delay. Any delay in the critical path can significantly impact the project timeline and budget. In the case of the development project proposal, the critical path was identified with a total duration of 50 weeks.

To mitigate this risk, the project team can optimize activities along the critical path and allocate sufficient resources to complete them on time. Project schedules and time analysis using PERT analysis result in findings impacting data management and strategic planning. Accurate project scheduling and time management are critical to successful data management initiatives.

7.1.2 The business case for implementing a data governance platform

The proposed platform, which includes a user-friendly interface and the ability to migrate data from Excel, enables staff to manage and query data more efficiently, saving time and increasing productivity. The platform can also automatically update child documents when parent documents or parent databases change, thus reducing the risk of errors and misunderstandings.

The potential risks associated with implementing a data governance platform include data breaches, compliance issues, and employee resistance; however, a well-defined risk management strategy and careful planning can mitigate these risks. The benefits of implementing a data governance platform far outweigh the risks. The platform can lead to significant long-term cost savings, improved processes, and a better customer experience while increasing efficiency and productivity and reducing the risk of errors and misunderstandings. Data governance practices, such as reducing duplicate data, improving data quality, and ensuring compliance with regulations, can also indirectly reduce costs and increase efficiency.

While the initial costs of implementation and licenses may be high, they can be recouped within the first year, followed by consistent savings in subsequent years. Additionally, the platform sets a precedent for the rest of the organization, promoting improved data management across departments and potentially leading to improved business outcomes.

7.1.2.1 Data governance platform implementation

To ensure the success of the platform implementation, the project team should conduct a thorough analysis of the existing data management processes and identify areas for improvement. This includes assessing the data quality, the data sources' reliability, and the effectiveness of the current data management practices. The team should also identify the key stakeholders affected by the implementation of the data governance platform and engage with them throughout the project to ensure their buy-in and support.

The project plan is considered a base aspect of the implementation strategy. The team, especially the project manager, will be tasked to develop a detailed implementation plan outlining the steps in migrating data from the existing systems to the new platform. This includes identifying the data sources, cleaning and preparing the data for migration, and configuring the new platform to integrate with the existing systems. Among other things, the team should also develop and agree on a testing plan to ensure that the new platform functions correctly and that the migrated data is accurate and complete.

Implementing a data governance platform is a complex project that requires careful planning, execution, and management. To achieve that, a thorough analysis of existing data management practices, selecting an appropriate platform, developing a detailed implementation strategy, and providing comprehensive training and support to users should be conducted. The team should also develop a risk management plan, a contingency plan, and a plan for measuring the platform's success.

The non-IT department can implement a data governance platform that improves data quality, efficiency, and cost savings, leading to better decision-making and an improved customer experience. Even though it was established that the project would be costly and difficult, it was still concluded that implementing the proposed data governance platform in the non-IT department of a multinational corporation has the potential to significantly improve departmental performance, increase efficiency and productivity, and reduce the risk of errors and misunderstandings, ultimately leading to a better customer experience and improved business outcomes.

7.2 Interview findings conclusion

The interviews conducted with five different respondents from various companies and departments reveal several critical insights into the experiences and needs of potential users of a data management platform.

One important finding is that data accuracy and up-to-date information are significant user concerns. Many spend time manually updating data and dealing with outdated folders, such as the

Junior Analyst who uses comparison techniques and tools like Power Query in Excel. These users would benefit from a platform that ensures data accuracy and reduces the need for manual data management.

Another key finding is that users often encounter difficulties finding and accessing data across multiple repositories and platforms. For example, the Data Security Employee manually searches through repositories to find data, which is time-consuming. A data management platform that connects all departments and streamlines data retrieval and sharing would be highly beneficial.

Data security and role-based access control are essential aspects of a potential platform, as highlighted by the Decision Maker's strict policies and the Hospital Employee's concerns about data sharing. Different users have varying levels of technical ability and responsibility, and a platform that accommodates these differences would be valuable.

The interviews also reveal a desire for a platform that offers flexibility in updating reports, version history features, and real-time data updates. As the Decision Maker and the Data Analyst suggested, these features could significantly improve data management processes and increase efficiency. Additionally, users expressed interest in integrating existing data sources and workflows, such as digitizing old paper records mentioned by the Hospital Employee, to create a more comprehensive and practical data management solution.

8 Conclusion

The thesis's main objective was to create a comprehensive business model for a data management platform. The importance and potential benefits of implementing a data management platform were emphasized throughout the research, highlighting its ability to significantly improve organizational performance and save companies money by increasing efficiency and reducing errors.

The Lean canvas approach was used to create the business model, which involved conducting market and strategic analyses using several frameworks such as SWOT, PEST, and Porter's five forces model. These analyses helped to identify potential opportunities and threats associated with implementing the platform, allowing the team to develop a risk management strategy. By analysing the market and strategic factors, the team was able to determine the feasibility of implementing the data management platform, identify potential challenges, and formulate a plan to mitigate these challenges.

To ensure that the project was completed in a timely and efficient, the PERT method was utilized to propose an estimate of the time required to develop the platform from scratch.

The research also involved interviewing potential customers for their opinions on the platform's benefits. The interviews were analysed using the design thinking method, allowing the author to identify the platform's unique features that would add users' value. Respondents described a unique platform feature that could differentiate it from the competition and provide added value. This feature was proposed as a wireframe in the last chapter, highlighting the potential for future innovation and development of the platform.

During the interviews, one of the respondents presented a potential business case. This business case contained several elements, including an executive summary, a mission statement, a product/service description, project organization details, financial appraisals, a project plan, identifying costs, a two-year cost plan, estimated savings, a return-on-investment analysis, a market assessment, and a risk assessment. These elements were thoroughly analysed to create a comprehensive risk management strategy to mitigate potential risks associated with implementing a data management platform.

The financial evaluation of the business case concluded that implementing a data management platform in an organization could lead to significant cost savings and process improvements. The initial costs for implementation and licenses are expected to be high but can be recouped within the first year, followed by consistent savings in subsequent years. The return-on-investment analysis showed that implementing a data management platform could lead to positive returns in the long term, highlighting the importance and potential benefits of such a platform in an organization.

8.1.1 Recommendations for further research

Further research could include analysing the potential impact of integrating artificial intelligence or machine learning algorithms into the platform to improve its capabilities further. This could allow the platform to provide more advanced and accurate data analysis, improving the overall effectiveness of the platform. Exploring potential customer acquisition strategies and analysing the possibility of extending the platform's services. The researcher can identify the best way to reach potential users and effectively market the platform by exploring potential customer acquisition strategies.

References

Agarwal, Preeti, P M Kothari Gautam Jain Yu Chen Hu Preeti Bajaj John Chibya Mubya P Malyadri Jerry Luftman Kirti Mahajan Prashant Varke P T Choudhari Parag Narkhede R H Gupta, R B Waghulade Rahul Kulkarni B H Barhate BJ Lathi Yogesh Mahajan S N Khan, and Sanjay Sugandhi. 2009. "Strategic Formulation Using Tows Matrix – A Case Study." In *Atharva Publications International Journal of Research and Development International Journal of Research and Development*, 1:87–90. www.ijrdonline.com.

Alex Osterwalder, Yves Pigneur, Greg Bernarda, and Alan Smith. 2014. "Value Proposition Design." *John Wiley & Sons, Inc., Hoboken, New Jersey*.

Alhassan, Ibrahim, David Sammon, and Mary Daly. 2016. "Data Governance Activities: An Analysis of the Literature." *Journal of Decision Systems* 25. https://doi.org/10.1080/12460125.2016.1187397.

Al-Ruithe, Majid, Elhadj Benkhelifa, and Khawar Hameed. 2019. "A Systematic Literature Review of Data Governance and Cloud Data Governance." *Personal and Ubiquitous Computing* 23 (5–6). https://doi.org/10.1007/s00779-017-1104-3.

"Alteryx Analytics Automation Platform | Alteryx." n.d. Accessed March 14, 2023. https://www.alteryx.com/alteryx-analytics-automation-platform.

Baits, Hamzah Abdul, Ika Arum Puspita, and Achmad Fuad Bay. 2020. "Combination of Program Evaluation and Review Technique (PERT) and Critical Path Method (CPM) for Project Schedule Development." *International Journal of Integrated Engineering* 12 (3). https://doi.org/10.30880/ijie.2020.12.03.009.

Bellman, Richard, Charles E Clark, Donald G Malcolm, Clifford J Craft, and Franc M Ricciardi. 1957. "On the Construction of a Multi-Stage, Multi-Person Business Game." *Source: Operations Research* 5 (4): 469–503.

Benzaghta, Mostafa Ali, Abdulaziz Elwalda, Mousa Mousa, Ismail Erkan, and Mushfiqur Rahman. 2021. "SWOT Analysis Applications: An Integrative Literature Review." *Journal of Global Business Insights* 6 (1). https://doi.org/10.5038/2640-6489.6.1.1148.

Camacho, Maria. 2016. "David Kelley: From Design to Design Thinking at Stanford and IDEO." https://doi.org/10.1016/j.sheji.2016.01.009.

Casadesus-Masanell, Ramon, and Joan Enric Ricart. 2010. "From Strategy to Business Models and onto Tactics." *Long Range Planning* 43 (2–3): 195–215. https://doi.org/10.1016/J.LRP.2010.01.004.

Ching, Hong Y. 2013. "Criticisms, Variations and Experiences With Business Model Canvas." *European Journal of Agriculture and Forestry Research* 1 (2): 26–37.

Cichy, Corinna, and Stefan Rass. 2019. "An Overview of Data Quality Frameworks." *IEEE Access* 7. https://doi.org/10.1109/ACCESS.2019.2899751.

"Collibra Doubles Valuation as Investors Chase next Cloud Winner." n.d. Accessed February 12, 2023. https://www.cnbc.com/2021/11/09/collibra-doubles-valuation-as-investors-chase-next-cloud-winner.html.

"Czech Industry Analysis 2022: Smart Industry 4.0 Technologies Help in the Crisis - News Service - Czech Technical University in Prague." n.d. Accessed February 11, 2023. https://aktualne.cvut.cz/en/reports/20220606-czech-industry-analysis-2022-smart-industry-40-technologies-help-in-the-crisis.

"Czech Republic Projection Note OECD Economic Outlook November 2022 by OECD - Issuu." n.d. Accessed February 11, 2023. https://issuu.com/oecd.publishing/docs/czech-republic-oecd-economic-outlook-projection-no.

"Czech Republic: Staff Concluding Statement of the 2022 Article IV Mission." n.d. Accessed February 11, 2023. https://www.imf.org/en/News/Articles/2022/11/22/mcs112222-czech-republic-staff-concluding-statement-of-the-2022-article-iv-mission.

"Czech Republic Tackles Gender Digital Divide as Part of EU Presidency - Kafkadesk." n.d. Accessed February 11, 2023. https://kafkadesk.org/2022/11/14/czech-republic-tackles-gender-digital-divide-as-part-of-eu-presidency/.

DaSilva, Carlos M., and Peter Trkman. 2014. "Business Model: What It Is and What It Is Not." *Long Range Planning* 47 (6): 379–89. https://doi.org/10.1016/j.lrp.2013.08.004.

"Data Catalog Market Analysis - Industry Report - Trends, Size & Share." n.d. Accessed February 12, 2023. https://www.mordorintelligence.com/industry-reports/data-catalog-market.

"Data Governance | Collibra." n.d. Accessed March 14, 2023. https://www.collibra.com/us/en/products/data-governance.

"Data Governance Market Analysis, Size And Trends Global Forecast To 2022-2030." n.d. Accessed February 9, 2023. https://www.thebusinessresearchcompany.com/press-release/data-governance-market-2022.

"Data Governance Tools and Products | Informatica." n.d. Accessed March 14, 2023. https://www.informatica.com/products/data-governance.html.

De, Henk, Swaan Arons, and Philip Waalewijn. 1999. "A Knowledge Base Representing Porter's Five Forces Model."

doc.Ing. Tomáš Šubrt, Ph.D., and Ph.D. Ing.Pavlína Langrová. 2013. *Projektové Řízení I*. Prague: Česká Zemědělská Univerzita v Praze.

Dorst, Kees. 2011. "The Core of 'design Thinking' and Its Application." *Design Studies* 32 (6): 521–32. https://doi.org/10.1016/J.DESTUD.2011.07.006.

Ehtisham Zaidi, Guido De Simoni. 2019. "Augmented Data Catalogs."

"Enterprise Data Catalog Software | Alation." n.d. Accessed March 14, 2023. https://www.alation.com/product/data-catalog/.

"EU - Political and Economic Environment." n.d. Accessed February 11, 2023. https://www.trade.gov/country-commercial-guides/eu-political-and-economic-environment.

Geissdoerfer, Martin, Marina P.P. Pieroni, Daniela C.A. Pigosso, and Khaled Soufani. 2020. "Circular Business Models: A Review." *Journal of Cleaner Production* 277 (December). https://doi.org/10.1016/J.JCLEPRO.2020.123741.

Greenberg, Jane. 2017. "Big Metadata, Smart Metadata, And." *Journal of Data and Information Science* 2 (3): 19–36. https://doi.org/10.1515/jdis-2017-0012.

Hooftman, Nils, Maarten Messagie, Joeri Van Mierlo, and Thierry Coosemans. 2018. "A Review of the European Passenger Car Regulations – Real Driving Emissions vs Local Air Quality." *Renewable and Sustainable Energy Reviews* 86 (April): 1–21. https://doi.org/10.1016/J.RSER.2018.01.012.

"How Russia Used SolarWinds To Hack Microsoft, Intel, Pentagon, Other Networks : NPR." n.d. Accessed February 12, 2023. https://www.npr.org/2021/04/16/985439655/a-worst-nightmare-cyberattack-the-untold-story-of-the-solarwinds-hack.

"IBM InfoSphere Information Governance Catalog | IBM." n.d. Accessed March 14, 2023. https://www.ibm.com/products/information-governance-catalog.

Jeroen Kraaijenbrink. 2022. "Updating the Business Model Canvas: 'The Strategy Sketch.'" February 19, 2022. https://www.jeroenkraaijenbrink.com/articles/strategy-sketch-business-model-canvas.

Joan Magretta. 2002. "Why Business Models Matter." *Harvard Business Review*, May.

Jones, Gardner M. 1960. “EDUCATORS, ELECTRONS, AND BUSINESS MODELS: A PROBLEM IN SYNTHESIS.: EBSCOhost.” 1960. https://web.s.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=d8bba3d2-118e-4098-b342-a022a36560dd%40redis.

Leifer, Larry, and Christoph Meinel. 2019. “Looking Further: Design Thinking Beyond Solution-Fixation.” In . https://doi.org/10.1007/978-3-319-97082-0_1.

Maurya Ash. 2022. *Running Lean: Iterate from Plan A to a Plan That Works. O’Reilly Media, Inc.* 3rd Editio. O’Reilly Media, Inc.

McNamara, C. 2009. “General Guidelines for Conducting Interviews. Retrieved January 11, 2010.” *Free Management Library*.

Nielsen, Christian, and Morten Lund. 2018. “A Brief History of the Business Model Concept.” *SSRN Electronic Journal*, no. March 2015. https://doi.org/10.2139/ssrn.2579439.

Osterwalder, Alexander, and Yves Pigneur. 2010. “Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers | Enhanced Reader.” John Wiley & Sons, Inc., Hoboken, New Jersey. 2010.

Osterwalder, Alexander, Yves Pigneur, Christopher L Tucci, Alexander ; Osterwalder, and Yves ; Pigneur. 2005. “Clarifying Business Models: Origins, Present, and Future of the Concept.” *Communications of the Association for Information Systems* 16: 1–25. https://doi.org/10.17705/1CAIS.01601.

Otto, Boris. 2011. “Data Governance.” *Business and Information Systems Engineering* 3 (4): 241–44. https://doi.org/10.1007/s12599-011-0162-8.

“Petr Pavel Elected President of the Czech Republic - Prague, Czech Republic.” n.d. Accessed February 12, 2023. https://www.expats.cz/czech-news/article/petr-pavel-elected-president-of-the-czech-republic.

Rahul, Kumar, and Rohitash Kumar Banyal. 2020. “Data Life Cycle Management in Big Data Analytics.” In *Procedia Computer Science*. Vol. 173. https://doi.org/10.1016/j.procs.2020.06.042.

Razzouk, Rim, and Valerie Shute. 2012. “What Is Design Thinking and Why Is It Important?” *Review of Educational Research* 82 (3): 330–48. https://doi.org/10.3102/0034654312457429.

Rob, Peter, Carlos Coronel, and Keeley Crockett. 2019. *Database Systems: Design, Implementation & Management*. *Nippon Ronen Igakkai Zasshi. Japanese Journal of Geriatrics*. Vol. 56.

Rocha, Arturo Orea, Carmelo García Pérez, Fernando Crecente Romero, and Maria Teresa Del Val Núñez. 2018. “The Business Model and Its Core Elements. Proposal of Definition and Table of Core Elements.” *Contemporary Economics* 12 (4 Special Issue): 497–518. https://doi.org/10.5709/ce.1897-9254.293.

Schmieden, Karen von, Lena Mayer, Mana Taheri, and Christoph Meinel. 2019. “An Iterative Approach to Online Course Design: Improving a Design Research MOOC.” In . https://doi.org/10.1007/978-3-319-97082-0_6.

Shabanova, L. B., G. N. Ismagilova, L. N. Salimov, and M. G. Akhmadeev. 2015. “PEST - Analysis and SWOT - Analysis as the Most Important Tools to Strengthen the Competitive Advantages of Commercial Enterprises.” *Mediterranean Journal of Social Sciences* 6 (3). https://doi.org/10.5901/mjss.2015.v6n3p705.

Stiftung, Bertelsmann. n.d. “Czech Republic.” Accessed February 11, 2023. https://www.bti-project.org.

Susan Moore. 2018. “How To Create A Business Case For Data Quality Improvement.” 2018. https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement.

Taipale-Erävala, Kyllikki, Erno Salmela, and Hannele Lampela. 2020. “Towards a New Business Model Canvas for Platform Businesses in Two-Sided Markets.” *Journal of Business Models*. Vol. 8.

Talend (Nasdaq: TLND). 2021. “THREE DATA GOVERNANCE SUCCESS STORIES: SECURE DATA, REDUCED COSTS, AND BETTER CUSTOMER EXPERIENCES.” 400 South El Camino Real, Suite 1400 . https://www.talend.com/resources/data-governance-case-studies/?ty=content.

Tanya Sammut-Bonnici and David Galea. 2014. “‘PEST Analysis’ in: Wiley Encyclopedia of Management Online | Enhanced Reader.” 2014.

Teece, David J. 2010. “Business Models, Business Strategy and Innovation.” *Long Range Planning* 43 (2–3): 172–94. https://doi.org/10.1016/J.LRP.2009.07.003.

Turner, Daniel W. 2010. “Qualitative Interview Design: A Practical Guide for Novice Investigators.” *Qualitative Report* 15 (3): 754–60. https://doi.org/10.46743/2160-3715/2010.1178.

“What Is Machine Learning Data Catalogs? Benefits and Use-Cases.” n.d. Accessed February 12, 2023. https://www.expressanalytics.com/blog/heres-why-machine-learning-data-catalogs-are-becoming-popular/.

Wirtz, Bernd W., Adriano Pistoia, Sebastian Ullrich, and Vincent Göttel. 2016. “Business Models: Origin, Development and Future Research Perspectives.” *Long Range Planning* 49 (1): 36–54. https://doi.org/10.1016/J.LRP.2015.04.001.

9 List of pictures, tables, graphs and abbreviations

Appendix

9.1 Interview structure and interview questions

Introduction (5 minutes)

a. Briefly explain the purpose of the interview

b. Explain the structure of the interview and confirm the interviewee's availability

c. Clarify that the interview is confidential and anonymous

Current Process for Working with Data (10 minutes)

a. Ask questions 2-4 to understand the interviewee's current process for working with data

Data Accuracy and Management (10 minutes)

a. Ask questions 5-11 to understand the interviewee's experience with data accuracy and management

V. Interest in a Data Management Tool (10 minutes)

a. Ask questions 12-16 to understand the interviewee's interest in a data management tool and their vision for its integration into their work process

Challenges with Data Management (5 minutes)

a. Ask question 17 to understand the interviewee's perspective on the challenges with data management in the company

Conclusion (5 minutes)

a. Summarize the key takeaways from the interview

b. Thank the interviewee for their time and participation

1. Can you describe your current process for working with data in your company?
2. How do you ensure the data you're using is accurate and up-to-date?
3. Have you ever faced a situation where you couldn't find the data you were looking for? How did you handle it?
4. What would you do if you stumbled upon an outdated document?
5. How do you keep track of the changes made to a document over time?
6. Can you imagine a scenario where a document could automatically update its information? How would that impact your work process?
7. Imagine that you are designing this feature. How do you imagine it would look like?
8. How important is it for you to clearly understand the lineage of the data you use?
9. Have you encountered any challenges with data governance in your company?

10. Would you like a platform that helps you manage data governance more efficiently?
11. How do you envision the integration of such a platform in your current work process?
12. Would you be interested in a tool that could help you manage data more efficiently and effectively?
13. How do you imagine such a tool fitting into your current work process?
14. Have you used any tools for data management in the past? If so, what was your experience like?
15. How do you envision your team benefitting from a tool that facilitates data management and governance?
16. How do you keep track of different versions of data and prevent data duplication?
17. In your opinion, what are some of the biggest challenges with data management in your company?

PERT - calculate the activity start and end times and their variances

To calculate the activity start and end times and their variances, we can use the following formulas:
Ti (0) = earliest start time for activity i
Ti (1) = earliest finish time for activity i
Tj (0) = earliest start time for activity j (successor of i)
Tj (1) = earliest finish time for activity j (successor of i)
σ2 (Ti (0)) = variance of the earliest start time for activity i
σ2 (Ti (1)) = variance of the earliest finish time for activity i
σ2 (Tj (0)) = variance of the earliest start time for activity j (successor of i)
σ2 (Tj (1)) = variance of the earliest finish time for activity j (successor of i)

The calculations involve the forward and backward passes in the network diagram. In the forward pass, we start with the first activity (A) and calculate its earliest start time (Ti (0)) and earliest finish time (Ti (1)) based on the optimistic (aij), most likely (mij), and pessimistic (bij) estimates. The formulas used for the forward pass are:

Ti (0) = 0 (for the first activity)

Ti (1) = Ti (0) + (aij + 4mij + bij)/6

For each succeeding activity, we calculate its earliest start and finish times based on the earliest finish time of the preceding activity (Ti (1)) and the duration of the activity. The formulas used for the forward pass are:

Ti (0) = max (Ti (1) of all preceding activities)

Ti (1) = Ti (0) + (aij + 4mij + bij)/6

In the backward pass, we start with the last activity (T) and calculate its latest start time (Ti (0)) and latest finish time (Ti (1)) based on the earliest finish time of the last activity and the duration of the activity. The formulas used for the backward pass are:

Ti (1) = T (1)

Ti (0) = Ti (1) – (aij + 4mij + bij)/6

For each preceding activity, we calculate its latest start and finish times based on the latest start time of the succeeding activity (Ti (0)) and the duration of the activity. The formulas used for the backward pass are:

Ti (1) = min (Ti (0) of all succeeding activities)

Ti (0) = Ti (1) – (aij + 4mij + bij)/6

The variances for each activity start and end time can be calculated as follows:

σ2 (Ti (0)) = sum of variances of earliest start times of all preceding activities

σ2 (Ti (1)) = sum of variances of earliest finish times of all preceding activities

σ2 (Tj (0)) = sum of variances of earliest start times of all succeeding activities

σ2 (Tj (1)) = sum of variances of earliest finish times of all succeeding activities

The variances can be calculated using the formula:

σ2 = [(bij – aij)/6]2

(doc.Ing. Tomáš Šubrt and Ing.Pavlína Langrová 2013)

Activity ID	Optimistic	Most Likely	Pessimistic	Duration (Te)	Standard deviation (σ)	Variance (σ^2)	Predecessor Activity/ies	Ti (0)	Ti (1)	Tj (0)	Tj (1)	σ^2 (Ti (0))	σ^2 (Ti (1))	σ^2 (Tj (0))	σ^2 (Tj (1))
1	1	4	7	4.33	1.16	1.35	None	0	0	4.33	4.33	0	0.67	0.67	0
1.1	1	2	3	2	0.5	0.25	1	0	0	2	2	0	0.25	0.25	0
1.2	1	4	7	4	1	1	1.4	0	0	4	4	0	1	1	0
1.3	2	3	4	3	0.33	0.11	1	0	0	2	2	0	0.11	0.11	0
1.4	2	4	6	4	0.67	0.44	1.3, 1.1	2	4	4	6	0.44	0.44	0.44	0.44
2.1	1	3	5	3	1.33	1.78	1.4	4	6	7	9	1.78	1.78	2.56	2.56
2.2	2	4	6	4	0.67	0.44	1, 2.4	4	8	4	8	0.44	0.44	0.44	
2.3	2	3	4	3	0.33	0.11	2.1	2	3	4	5	0.11	0.11	0.11	0.11
2.4	2	4	6	4	0.67	0.44	1.2, 2.2	3	4	6	7	0.44	0.44	0.44	0.44

3.1	3	6	9	6	1.33	1.78	2.4	3	6	9	10	1.78	1.78	1.78	1.78
3.2	4	7	10	7	1.67	2.78	3.1, 3.3	7	8	17	18	7.44	7.44	7.44	7.44
3.3	2	5	8	5	1.33	1.78	2.3	5	8	13	16	1.78	1.78	1.78	1.78
3.4	3	6	9	6	1.33	1.78	3.2	6	9	15	18	1.78	1.78	1.78	1.78
4.1	5	8	11	8	1.33	1.78	3.4	8	11	23	26	1.78	1.78	1.78	1.78
4.2	2	5	8	5	1.33	1.78	4.1	5	8	13	16	1.78	1.78	1.78	1.78
4.3	2	4	6	4	0.67	0.44	4.1	5	6	11	12	0.44	0.44	0.44	0.44
4.4	3	5	7	5	0.67	0.44	4.2, 4.3	6	8	14	16	0.44	0.44	0.44	0.44
5.1	3	5	7	5	0.67	0.44	4.4	6	8	14	16	0.44	0.44	0.44	0.44
5.2	3	6	9	6	1.33	1.78	4.4	9	11	20	22	1.78	1.78	1.78	1.78
5.3	3	5	7	5	0.67	0.44	4.4	6	8	14	16	0.44	0.44	0.44	0.44

5.4	2	4	6	4	0.67	0.44	4.4	4	6	10	12	0.44	0.44	0.44	0.44
6	2	4	7	4.33	0.94	0.89	5.1, 5.2, 5.3, 5.4	2	4	7	4.33	0.94	0.89	6.12	1.35

Table 20PERT Project Schedule and Time Analysis

The table shows the calculation of activity start and finish times and variances for a project with multiple activities. Activities are listed from first to last.

For each activity, four columns show the start and end times of the activity, as well as the deviations associated with these times. The start and end times are given in activity durations, with the start time denoted as Ti (0) and the end time as Ti (1). Also listed are the next activity's corresponding start and end times denoted as Tj (0) and Tj (1).

The last four columns show each activity's start and finish times variances. These variances are used to calculate the standard deviation of each activity, which in turn is used to calculate the probability of completing the project within a specific time frame.

The variances for each activity are determined based on the PERT formula, which considers the optimistic, most likely, and pessimistic estimates of the duration of the activity. A minor deviation means a more accurate estimate of duration, while a more significant deviation means more uncertainty. (doc.Ing. Tomáš Šubrt and Ing. Pavlína Langrová 2013; Baits, Puspita, and Bay 2020)

The table below presents the Work Breakdown Structure (WBS) for the Data governance platform project, broken down by activity phase, activity ID, description and duration. Estimates are in the form of optimistic, most likely, and pessimistic scenarios, and time (Te), Standard Deviation (σ), and variance (σ^2) are calculated based on the Program Evaluation and Review Technique (PERT). The Prior Activity (s) column identifies activities to be completed before the current activity. The work breakdown structure is a helpful tool for project planning and management, allowing for better control and transparency of project progress.

Table 21Business case activities duration

9.1.1 Business Case time analysis

Stage	Activity ID	Activity Description	Optimistic	Most Likely	Pessimistic	Duration (Te)	Standard Dev (σ)	Variance (σ^2)	Predecessor Activity/ies
Project Initiation	1.1	Define project scope, objectives, and success criteria	1	2	3	2	0.33	0.11	N/A
	1.2	They form a project team and allocate roles and responsibilities	1	2	3	2	0.33	0.11	1.1
	1.3	Conduct stakeholder analysis and engagement planning	1	2	3	2	0.33	0.11	1.1
	1.4	Define project timeline and budget	1	2	3	2	0.33	0.11	1.1, 1.3
Requirements Gathering	2.1	Identify and document requirements for the platform	2	4	6	4	1	1	1.2
	2.2	User research to understand user needs and workflows	2	4	6	4	1	1	1.2
	2.3	Define data governance policies and procedures	2	4	6	4	1	1	2.1, 2.2
Vendor Selection	3.1	Conduct market research to identify potential vendors	1	2	3	2	0.33	0.11	2.3
	3.2	Develop a request for proposal (RFP)	1	2	3	2	0.33	0.11	3.1
	3.3	Evaluate vendor proposals and select the preferred vendor	2	4	6	4	1	1	3.2
	3.4	Negotiate a contract with the selected vendor	2	4	6	4	1	1	3.3
Implementation	4.1	Install and configure the data governance platform	4	6	8	6	1	1	3.4
	4.2	Develop and execute a data migration plan	3	5	7	5	1.33	1.78	4.1
	4.3	Develop and execute a user training plan	2	4	6	4	0.67	0.44	4.1
	4.4	Conduct user acceptance testing and resolve any issues	3	4	5	4	0.33	0.11	4.2, 4.3
PostImplementation	5.1	Conduct post-implementation review	2	3	4	3	0.33	0.11	4.4
	5.2	Identify and resolve any issues	3	6	9	6	1.33	1.78	5.1
	5.3	Develop and implement maintenance	4	6	8	6	0.67	0.44	5.1
	5.4	Plan for future upgrades and enhancements	3	5	7	5	0.67	0.44	5.1
End of Project	6	Project closeout and documentation	2	3	5	3	0.67	0.44	5.2, 5.3, 5.4

This table provides the estimated duration, Standard Deviation, and variance for each activity in the project along with its predecessors. The Ti (0) and Ti (1) columns show the earliest and latest start times, respectively, for each activity, while the Tj (0) and Tj (1) columns show the earliest and latest finish times, respectively. The last four columns show the variance for each of these time estimates. This table provides a comprehensive overview of the project activities, their durations, and the critical path, which can be used to help plan and manage the project effectively.

Activity ID	Optimistic	Most Likely	Pessimistic	Duration (Te)	Standard Deviation (σ)	Variance (σ^2)	Predecessor Activity/ies	Ti (0)	Ti (1)	Tj (0)	Tj (1)	σ^2 (Ti (0))	σ^2 (Ti (1))	σ^2 (Tj (0))	σ^2 (Tj (1))
1	1	4	7	4.33	1.16	1.35	None	0	0	4.33	4.33	0	0.67	0.67	0
1.1	1	2	3	2	0.5	0.25	1	0	0	2	2	0	0.25	0.25	0
1.2	1	4	7	4	1	1	1.1	0	0	4	4	0	1	1	0
1.3	1	2	3	2	0.33	0.11	1.1	0	0	2	2	0	0.11	0.11	0
1.4	2	4	6	4	0.67	0.44	1.1, 1.3	2	4	4	6	0.44	0.44	0.44	0.44
2.1	2	4	6	4	1	1	1.2	4	6	7	9	1.78	1.78	2.56	2.56
2.2	2	4	6	4	0.67	0.44	1.2, 2.4	4	8	4	8	0.44	0.44	0.44	0.44
2.3	2	4	6	4	1	1	2.1, 2.2	5	8	4	7	1	1	1	1
2.4	2	4	6	4	0.67	0.44	1.2, 2.2	3	4	6	7	0.44	0.44	0.44	0.44
3.1	3	6	9	6	1.33	1.78	2.3	3	6	9	10	1.78	1.78	1.78	1.78
3.2	4	7	10	7	1.67	2.78	3.1, 3.3	7	8	17	18	7.44	7.44	7.44	7.44
3.3	2	5	8	5	1.33	1.78	2.3	5	8	13	16	1.78	1.78	1.78	1.78
3.4	2	4	6	4	1	1	3.3	6	8	14	16	1	1	1	1
4.1	5	8	11	8	1.33	1.78	3.4	8	11	23	26	1.78	1.78	1.78	1.78
4.2	3	5	7	5	1.33	1.78	4.1	6	8	11	13	1.78	1.78	1.78	1.78
4.3	2	4	6	4	0.67	0.44	4.1	6	7	10	11	0.44	0.44	0.44	0.44
4.4	3	4	5	4	0.33	0.11	4.2, 4.3	9	11	15	17	0.11	0.11	0.11	0.11
5.1	2	3	4	3	0.33	0.11	4.4	13	14	19	20	0.11	0.11	0.11	0.11
5.2	3	6	9	6	1.33	1.78	4.4	16	20	28	32	1.78	1.78	1.78	1.78

5.3	4	6	8	6	0.67	0.44	4.4	14	16	22	24	0.44	0.44	0.44	0.44
5.4	3	5	7	5	0.67	0.44	4.4	11	13	18	20	0.44	0.44	0.44	0.44
6	2	3	5	3	0.67	0.44	5.1, 5.2, 5.3, 5.4	20	23	38	41	0.44	0.44	0.44	0.44

Table 22 Business Case PERT