MASTER

A combined Machine Learning and DMN Approach for Knowledge-Intensive Decision-Making within Open Innovation
An Explorative Case Study

Reuvekamp, M.C.

Award date:
2022

Eindhoven University of Technology
Department of Industrial Engineering and Innovation Sciences

A combined Machine Learning and DMN Approach for Knowledge-Intensive Decision-Making within Open Innovation

An Explorative Case Study

Master Thesis in partial fulfillment of the requirements for the degree of
Master of Science in Operations Management and Logistics

Author:
M.C. (Manon) Reuvekamp – 1369059

Supervisors/assessors:
dr. ir. H. Eshuis – TU/e, Information Systems
dr. B. E. Aysolmaz – TU/e, Information Systems
dr. I. D. C. Grau – TU/e, Information Systems
dr. F. Geuzebroek – Shell, Senior Science Expert

July 4, 2022

“In the fast-changing business world of today, innovation has become the mainstay of organizations. The complexity of innovation has been increased by growth in the amount of knowledge available to organizations (Du Plessis, 2007).”

Abstract

This explorative case study focuses on improving decision-making capability within knowledge-intensive open innovation processes by deriving decision logic at decision points using historical data. For this study, a case at Shell was used, which was conducted at one of their open innovation processes. This study aims to identify the first steps toward an approach that improves decision-making capability (the ability for supporting decisions based on historical insights) within knowledge-intensive innovation processes through the case at Shell. To do so, this study proposes a Decision Tree (Classifier) as an ML model to derive decision logic at decision points via decision-mining and Decision Model and Notation (DMN) to represent the mined decision logic. The combined approach results in a supportive DMN model that improves process transparency and relieves Shell's experts by a decision model that filters, i.e. separates, the (un)successful open innovation projects based on the project's characteristics. Both the ML model's 80% accuracy and well-scored experts' validation confirm that the approach properly utilizes historical data to improve decision-making capability. Under the conditions of pre-screening the dataset and good data quality as input, the combined ML and DMN approach can be generalized to various contexts with the assumptions and resources committed. *Note: some company data has been censored and/or replaced by substitutes such as [A] or [1].*

Keywords: Knowledge-Intensive (Processes), Decision-Making (Capability), Open Innovation, Decision Model and Notation (DMN), Decision Trees, Machine Learning

Summary
This explorative case study focuses on improving decision-making capability within knowledge-intensive open innovation processes by deriving decision logic at decision points using historical data. For this study, a business case at Shell was used which was conducted at one of Shell's open innovation processes called the GameChanger (GC) process. Case study research methodology is implemented to generalize the context-specific findings of Shell's case to various contexts.

Problem context
Currently, the decision-making within Shell's open innovation process fully relies on expert knowledge that is based on an expert's past experience. It is unknown what the decision-making capability is and historical data is not utilized while this valuable historical data is available. As a result, Shell lacks insights on the criteria that they can deduce from their past projects to improve decision-making in the open innovation process. Therefore, the problem statement of this study for the business case is:

The current open innovation process at Shell does not utilize historical data to gain insights into decision-making capability

Supporting decision-making within open innovation is critical for investigating as early as possible whether an open innovation project should be decided to select or not. As open innovation projects are of high strategic importance where many resources are involved, resources should be allocated to the most valuable projects. Therefore, it is critical to make a *GO* or *NO-GO* decision as early as possible once a project is selected to allocate resources best to the projects to be most likely successful. However, such insights are difficult to gather in open innovation because these processes are based on knowledge and ideas. Moreover, organizations require better approaches/techniques to derive such insights from their historical data as this valuable data might not be (properly) utilized currently. Although some characteristics of open innovation equal characteristics of KiPs such as its uncertainty, complexity, and knowledge intensity (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), open innovation differs from usual KiPs. Open innovation is specifically of high strategic importance (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), particular risky (Banu et al., 2016; Herzog, 2008; Saura et al., 2022; Wu et al., 2013), and operates in highly dynamic environments (Banu et al., 2016). Therefore, they require specific approaches/techniques for deriving insights in decision-making using historical data.

The main objective of this explorative case study is to define an approach that improves decision-making capability, referring to the ability to support decisions based on historical insights (Ghattas et al., 2014), in open innovation processes through the business case at Shell. This study aims to identify the first steps towards a general solution to the problem context using the business case at Shell. To fill in the gap of missing insights and the need for well-fitting techniques for the specific context, this study proposes a combined approach of Decision Tree (Classifier) as Machine Learning (ML) model to derive decision logic at decision points via decision-mining and Decision Model and Notation (DMN) to represent the mined decision logic to suitably improve knowledge-intensive decision-making within open innovation. DMN is a presentation tool that fulfils the need for organizational decision-making and ensures flexibility, transparency, improved efficiency, and improved quality and compliance by standardizing processes and their decisions (Etinger et al., 2019).

Research questions

Based on the abovementioned problem context, the general main research question of this study is:

How can the decision-making within the open innovation process be supported by analysing historical data to improve decision-making capability?

To address this main research question, a case study at Shell was conducted. To be able to answer the main research question, the question is divided into four Shell-specific sub-questions in this case study. The first two sub-questions have a describing and exploring nature, the third sub-question is about designing the solution, and the fourth sub-question is about evaluating. The four sub-questions are:

1. *How is Shell's open innovation process currently designed?*
2. *What are the opportunities and limitations in Shell's open innovation process?*
3. *How can decision-making capability in Shell's open innovation process be improved?*
4. *How can the proposed solution design be evaluated?*

Methodology

Subsequently, the research methodology is based on case study research methodology as the study is an explorative case study. To execute the research, the case study design is structured based on the case study research design discussed by Yin (2009). The definition of the case is the preliminary stage of the case study research methodology, which is about formulating the theory, selecting the case, and defining research questions (Yin, 2009). As part of the fieldwork and analysis stage (Yin, 2009), the data collection (divided into a qualitative part and a quantitative part), and data analysis are described. To analyse the data and develop an approach for the case accordingly, a combination of two commonly used frameworks for problem-solving in industrial settings was used. These frameworks are well-known frameworks used to design improvement projects and help to analyse these projects in a problem-solving manner. The key methodology to implement the approach in the case is structured according to the Van Aken framework, which has a process-related perspective. Additionally, this study covers a data-related perspective, and this part is covered by the CRISP-DM framework. Figure 1 shows the abovementioned methodology used in this study. The mix of elements of the case study is shown in the middle of the figure as the core element. Steps 2-7 are used to answer the main research question (step 1) and to investigate to which extent the case study could be generalized to other contexts accordingly (step 8).

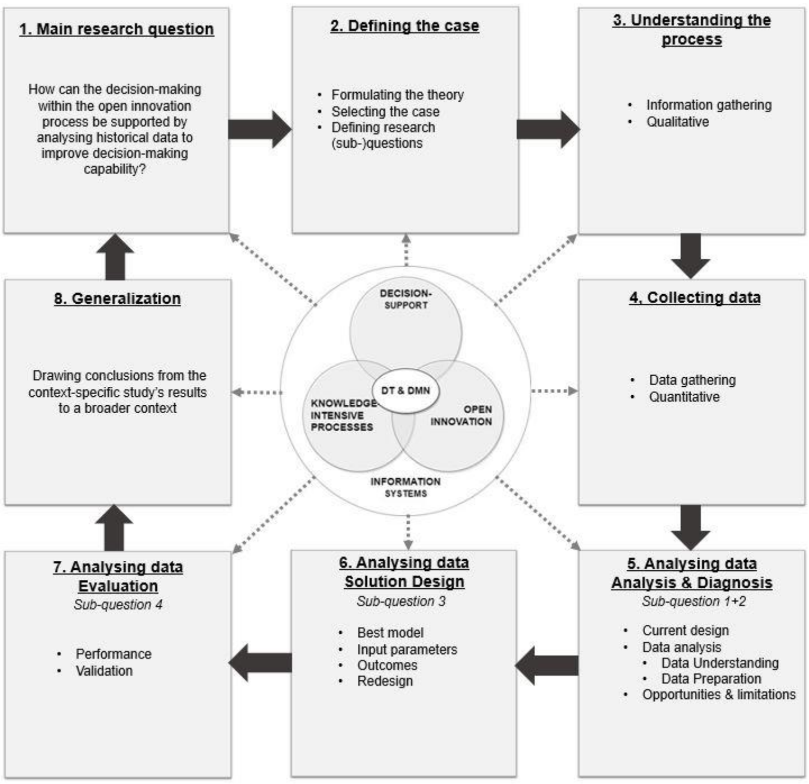


Figure 1. Case study design

Results

Currently, the open innovation process at Shell consists of multiple tasks, which together shape the current process flow. The most important distinction in the current design is the [M] phase versus the [N] phase, where the [M] phase considers the input for the decision model that serves as a filter for the expert knowledge model in the [N] phase. However, it is unknown what the decision-making capability is and historical data is not utilized while this valuable historical data is available for thorough analysis by an ML model. Therefore, a major opportunity arose to improve decision-making capability by deriving decision logic at decision points based on historical data. The investigated limitations are a prioritizing problem, deviated terminology used in the process flow, and a biased way of approaching the process flow. Finally, data-related challenges were uncovered where the most important one is inconsistent data registration relating to incomplete information. However, a major update in 2020 concerning data registration divides the dataset into two different datasets: the dataset *before GC update 2020* and the dataset *after GC update 2020*. The challenge of inconsistent data registration refers to the dataset *before GC update 2020* as this dataset has poor data quality while the dataset *after GC update 2020* is reliable, so the challenge does not hold for this dataset. The reason for this difference is that lacking poor data quality is solved in the newest dataset as since the update of 2020 most fields in the submission form are made mandatory. Therefore, the data consistency in the newest dataset is much better and has better data quality accordingly.

Ideally, both insights into characteristics belonging to successful projects and unsuccessful projects were derived from the ML model. However, deriving characteristics that belong to successful projects turned out difficult and the DT as ML model only gives insights into derived decision logic at decision points for the rejected class representing unsuccessful projects. However, these ML outcomes for the rejected class do give powerful insights and are supported by a significant number of projects. The biggest change in the redesigned process flow is the additional activity of making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. These insights improve the decision-making capability as the new approach improves process transparency and resources can be allocated to the projects to be likely successful. Additionally, Shell's experts are relieved by a decision model that filters, i.e. separate using the ML model, the successful and unsuccessful projects based on the project's characteristics. Although historical data is utilized and future decisions have better pillars, Shell's experts are still required as the decision model has a supportive nature. Based on the performances (especially the ML model's 80% accuracy) and experts' validation, it was concluded that the model's results are reliable and, therefore, the proposed solution design is realistic and feasible.

Conclusion

Finally, the decision-making within the open innovation process can be supported by analysing historical data to improve decision-making capability by the combined ML and DMN approach for knowledge-intensive decision-making within open innovation. Here, the DT as ML model identified decision logic at decision points via decision-mining and DMN represented the mined decision logic for visualization and interpretation purposes. The study's evaluation confirmed the reliable and powerful insights that this approach can achieve in the context of knowledge-intensive decision-making within open innovation. The most important lesson learned from the explorative case study is that utilizing historical data provides meaningful insights, which in turn ensures improved decision-making capability (referring to more informed decision-making) in organizations. Under the conditions of pre-screening the dataset and good data quality as input, the combined ML and DMN approach can be generalized to various contexts with the assumptions and resources committed.

Preface

This report represents my Master’s Thesis, which is the result of my graduation project and partly fulfils the requirements for the degree of a Master of Science in Operations Management and Logistics. The project was undertaken at the Eindhoven University of Technology in the Information Systems group in collaboration with Shell. This section allows me to thank a few people involved for their contribution to my project. First, many thanks to my supervisor at Shell, dr. F. Geuzebroek, for providing me with the opportunity to execute my graduation project at such an impressive organization. I also want to thank him for his availability, guidance, and feedback during the internship. During the internship’s first phase (proposal writing), we increasingly improved the project’s scope through extensive brainstorming sessions. Besides dr. F. Geuzebroek, I would like to thank my other colleagues at Shell for their cooperation and interesting points of view. I appreciate their efforts and time because without their cooperation I would not have been able to conduct the analyses.

Next, I would like to thank my first TU/e supervisor, dr. ir. H. Eshuis, for his extensive guidance and constructive feedback during this project. We frequently met each other to discuss the (progress of the) project and I appreciate his quick responses to any question or intermediate version. Furthermore, I would like to thank dr. B. E. Aysolmaz for being my second TU/e supervisor. She provided me with constructive feedback and valuable suggestions during the second phase (after proposal writing) of my graduation project. For me, aligning the practical perspectives and the theoretical perspectives was the most difficult part. Both supervisors guided me in the right direction and provided support to improve stepwise. Last, I would like to thank dr. I. D. C. Grau that she wanted to be my third assessor at short notice.

Finally, my appreciations and acknowledgements go to my family for their unconditional support during my studies, especially during this graduation project. This last phase of my studies to complete my Master of Science was a tough time due to multiple circumstances such as COVID-19 and personal circumstances. However, my family have supported and motivated me all the way, and I would like to express my gratitude to them. Despite the circumstances, I have always remained dedicated, and I enjoyed finally being able to put the theory into practice during this project. The graduation project was an interesting, challenging, and great (personal) learning experience.

Manon Reuvekamp
July 4, 2022

Table of Contents

List of Figures

List of Tables

List of abbreviations

BPM: Business Process Management

BPMN: Business Process Model and Notation

DMN: Decision Model and Notation

GC: Game Changer

IS: Information System(s)

KiP(s): Knowledge-intensive Process(es)

ML: Machine Learning

SLR: Structured Literature Review

TU/e: Eindhoven University of Technology

1 Introduction

Over the past decades, business processes have been changing in point of view regarding success. They have been changing from success in terms of predictability and efficiency to success in terms of capability to adapt and treat unique customer requests, which require creativity and collaboration and rely on the knowledge of experts (Boissier et al., 2019). Typically, these processes are knowledge-intensive processes (KiPs), which have several decision-making tasks and heavily rely on knowledge workers' expertise and experience (Venero et al., 2019). In addition, organizations are facing increasing complexity within their open innovation processes and are challenging decisions daily. Technology experts face many decisions within open innovation before being able to work on a new idea. Typically, knowledge is an unavoidable and fundamental resource for creating innovation and innovation is commonly the combination of knowledge and ideas (Wu et al., 2013).

1.1 Research motivation

Supporting decision-making within open innovation is critical for investigating as early as possible whether an open innovation project should be decided to select or not. As open innovation projects are of high strategic importance where many resources are involved, resources should be allocated to the most valuable projects. Therefore, it is critical to make a *GO* or *NO-GO* decision as early as possible once a project is selected to allocate resources best to the projects to be most likely successful. This is best achieved when decision-making is transparent and of good capability. However, such insights are difficult to gather in open innovation because these processes are based on knowledge and ideas. Moreover, organizations require better approaches/techniques to derive such insights from their historical data as this valuable data might not be (properly) utilized currently. Although some characteristics of open innovation equal characteristics of KiPs such as its uncertainty, complexity, and knowledge intensity (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), open innovation differs from usual KiPs. Open innovation is specifically of high strategic importance (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), particular risky (Banu et al., 2016; Herzog, 2008; Saura et al., 2022; Wu et al., 2013), and operates in highly dynamic environments (Banu et al., 2016). Therefore, they require specific approaches/techniques for deriving insights in decision-making using historical data. This explorative case study aims to use a Decision Tree (Classifier) as Machine Learning (ML) model to derive decision logic at decision points via decision-mining and Decision Model and Notation (DMN) to represent the mined decision logic to suitably improve knowledge-intensive decision-making within open innovation.

1.1.1 Business case problem

For the abovementioned research motivation of this study, a business case was used and this section explains the problem definition according to this industrial (or so-called business) case. Currently, the decision-making within Shell's open innovation process fully relies on expert knowledge that is based on an expert's past experience. As the open innovation process has multiple experts from multiple disciplines involved, it results in a biased way of approaching it due to the individuality (deviated interpretations and preferences) of multiple experts. There is no guidance for this decision-making that is based on patterns in historical data. Moreover, there is no documentation explaining what the exact process execution is and what the decision-making capability is, which refers to the ability to support decisions based on historical insights (Ghattas et al., 2014). Therefore, the problem statement of this study for the business case is: *the current open innovation process at Shell does not utilize historical data to gain insights into decision-making capability.*

Shell lacks insights on the criteria that they can deduce from their past projects to improve decision-making in the open innovation process. Right now, it is unknown which characteristics belong to successful projects and which characteristics belong to unsuccessful projects. Therefore, transparency of the process should be improved to improve decision-making capability. It is crucial to master this knowledge because it gives insights into possible success or failure criteria, which supports decisions in the future. Once these insights are clear, the greater the capability to as soon as possible reject an unsuccessful project or convert (proceed) a successful project.

The database of the Game Changer (GC) process deals with data generated during the operationalization and management of the process. This database contains detailed information about the characteristics of projects and the evaluations of decision-making over time. As of 2017, the GC's database offers consistent data, which is structured and could be of large queries depending on the filter area. The projects of this database are Shell-wide, which ensures that many people are working together, many alternatives are possible, and a high(er) budget is available compared to specific department projects. However, it is unknown what the decision-making capability is and the available historical data is not utilized. Therefore, a major opportunity arises to analyse historical data to derive decision logic at decision points to improve Shell's decision-making capability of the open innovation process.

1.1.2 Business case description

The abovementioned business case problem was conducted at Shell plc (former Royal Dutch Shell plc), known as Shell, which was founded in 1907 and comprises a global group of energy and petrochemical companies (Shell, 2022c). Shell is supported by more than 80,000 employees, located in more than 70 counties, and its revenue in 2021 was $261,504 million (Shell, 2022c). Shell formulates their strategy as 'powering progress', which accelerates the transition of its business to net-zero emissions in line with society accordingly (Shell, 2022b). This strategy ensures chasing the four main goals: generating shareholder value, achieving net-zero emissions, powering lives, and respecting nature (Shell, 2022b). As formulated by Shell, a set of core values are important within Shell's working environment, which are honesty, integrity and respect (Shell, 2022a). These core values are supported by manuals such as the Code of Conduct, which help Shell's employees to act in line with the core values defined (Shell, 2022a). The core activities of Shell are (Shell, 2022c):

> *"Using advanced technologies, adopting innovative approaches to help build a sustainable energy future, and investing in power (including wind and solar) and new fuels for transport (including advanced biofuels and hydrogen)"*

The operations of Shell are divided into multiple businesses, which are Upstream, Integrated Gas and Renewables and Energy Solutions, and Downstream (Shell, 2022b). In addition, the Projects & Technology organisation supports the delivery of major projects of Shell, which drives the research and innovation to support the development of new technology solutions (Shell, 2022b). The Research and Development (R&D) activities of Shell are carried out within the global network of technology centres within Projects & Technology (Shell, 2022b). Innovation and R&D are of huge importance for Shell and, therefore, they decided in 2016 to continue investing in their R&D; their R&D spending in 2016 was $1,014 million (Shell, 2017). The investment aimed to improve the efficiency of Shell's products, processes, and operations. Moreover, the investment focused on developing new technologies supporting the transition of the society to a low-carbon system (Shell, 2017). Besides in-house R&D activities, Shell complemented its R&D and innovation department through various collaborations with multiple universities and three different open innovation processes as described on its website (Shell, 2022d): Shell GameChanger (GC), Shell Technology Ventures (STV), and Shell TechWorks (STW). All these open innovation processes offer opportunities to other external parties to share both rewards and risks of innovation with Shell, which could be short-term technology developments as well as long(er)-term technology developments (Shell, 2022d).

This study's business case focuses on the GC process, which comprises fostering entrepreneurship by funding and supporting early-stage start-ups (Shell, 2022d). External parties can apply to the GC process by submitting an online submission form, and these parties can be industries and universities but are mostly companies. In more detail, the GC process provides financial and technical support to prove concepts of external parties that might apply to both the oil and gas sector and alternative energy courses. As of 1996, many innovates applied to the GC process, which may have the potential to become one of Shell's major projects (Shell, 2022d).

1.1.3 Business case scope

The responsible department for the business case is GC and, therefore, the GC process is the scoped process. The GC process deals with technology developments that commonly comprise a Technology Readiness Level (TRL) of approximately 2, 3, 4 or 5. TRL represents the depth in maturity and availability of the technology developments. According to Shell, all TRLs have an indicated formulation but these formulations are generalized. First, the definition of TRL 2 is technology concept and/or application formulated and 3 stands for analytical and experimental critical function and/or characteristic proof of concept. Lastly, TRL 4 comprises a component and/or validation in a laboratory environment whereas TRL 5 comprises a component and/or validation in a relevant environment.

The GC process consists of seven general stages, which are simplistically shown in Figure 2, but this study focuses on the process from submitting an idea via the online submission form that is connected to the GC portal until (i.e. not including) stage [5] as indicated by the black oval. Therefore, the scoped GC process within this study is: *the process of an innovative idea entering the system via an online submission form until (i.e. not including) reaching (or not) stage [5].* The first stage (submitting an idea) is widely known, but the rest of the process is replaced by stages [2]-[7] due to confidentiality.

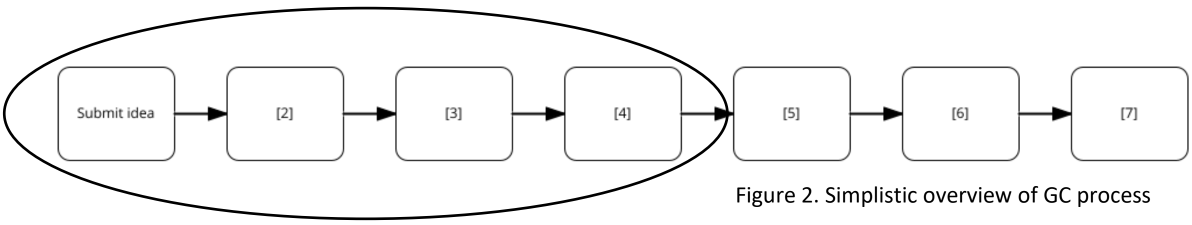


Figure 2. Simplistic overview of GC process

1.2 Research objective

The main objective of this explorative case study is to define an approach that improves decision-making capability, referring to the ability to support decisions based on historical insights (Ghattas et al., 2014), in open innovation processes. For this explorative case study, a business case at Shell was used which was conducted at the GC process representing one of Shell's open innovation processes. Case study research methodology is implemented to focus on generalizability and lessons learned from the context-specific findings of Shell's case in various contexts. This study aims to identify the first steps towards a general solution to the problem context using the business case at Shell. To do so, the research implication of this study is to provide a combined ML and DMN approach for knowledge-intensive decision-making within open innovation that can be used in various contexts. This approach improves decision-making capability as it derives decision logic at decision points based on an event log consisting of historical data.

Moreover, the study's practical implication is supporting Shell's experts with their decision-making and investigating the supportive nature of insights into decision-making capability based on historical data. Deriving decision logic at decision points that investigate which characteristics belong to successful projects and which characteristics belong to unsuccessful projects is the core element of this study. As a result, the greater the ability for Shell's experts to as soon as possible reject an unsuccessful project or convert (proceed) a successful project. In an ideal state, Shell wishes to only work on successful open innovation projects, because this will save a lot of resources (people, money, and time) as Shell's experts spend a lot of time working on these projects.

1.2.1 Research questions

Based on the earlier described sections, this section describes this study's main research question and its Shell-specific sub-questions for the case study. The general main research question of this study is:

How can the decision-making within the open innovation process be supported by analysing historical data to improve decision-making capability?

To address this main research question, a case study at Shell was conducted. To be able to answer the main research question, the question is divided into four Shell-specific sub-questions in this case study. These sub-questions are created using the Van Aken framework, which is widely used as a problem-solving methodology in organizations to design practical business improvement projects based on various aspects and steps (Van Aken et al., 2012). Van Aken et al. (2012) define a framework that considers five problem-solving phases: Problem Definition, Analysis & Diagnosis, Solution Design, Implementation, and Evaluation. There will not be an implementation phase within this study due to time issues and larger-scale research, and Chapter 3 explains in more detail the research methodology and the frameworks to implement the approach in the industrial case. The first two sub-questions have a describing and exploring nature, the third sub-question is about designing the solution, and the fourth sub-question is about evaluating. The four sub-questions of this research are:

1. *How is Shell's open innovation process currently designed?*
2. *What are the opportunities and limitations in Shell's open innovation process?*
3. *How can decision-making capability in Shell's open innovation process be improved?*
4. *How can the proposed solution design be evaluated?*

1.2.2 Research relevance

Besides the previously mentioned objective of this study, another objective is to contribute to academic knowledge. Therefore, providing related work as a base is crucial as this part discusses studies that are similar in content or goals. The Business Process Management (BPM) field fulfils the need for processes with an explicit focus on modelling and visualizing these processes. Instead of using a data-centric model such as CMMN or a routine model such as BPMN, it is proposed to use DMN modelling as decisions and expert knowledge could be perfectly shown. On the other hand, BPM is not commonly used for data analytics, however, Machine Learning (ML) is. ML is part of the well-known concept of Artificial Intelligence, and is about using computer algorithms and is of key importance in accessing data (Ali et al., 2022). As the growth of data is increasing daily, the amount of challenges in data is growing as well, which has led to the increasing need for emerging technologies rather than traditional methods (Lu, 2018). As a result, organizations are motivated to use upstream technologies such as ML, and Czvetkó et al. (2022) stated that the number of ML projects is expected to double within a year.

The combination of DMN with a Decision Tree (DT) is not unknown in scientific knowledge. Earlier research showed that DMN uses a decision logic level that can be created using a DT. For example, the paper of Etinger et al. (2019) presents a method for automatically creating DMN decision tables from a DT (Classifier). The benefit of using a DT is that this ML model is easily interpretable, however, the drawback is that a DT might lack accuracy compared to black-box models (Sachan et al., 2020). Therefore, instead of evaluating discovered decision rules on their accuracy, the paper of Scheibel & Rinderle-Ma (2021) focused on the semantics and meaningfulness of the results rather than accuracy. This implies that discovering decision rules is also not unknown in scientific knowledge and can be done using multiple evaluation ways to validate and implement the decision rules (Scheibel & Rinderle-Ma, 2021).

This research continues on the beforementioned work by building on existing approaches and evaluating these approaches in a business case specific to the open innovation environment. The research relevance from a scientific point of view is combining the Decision Tree model to derive decision logic at decision points via decision-mining combined with the quite new DMN modelling approach as a visualization tool for the mined decision logic in a knowledge-intensive open innovation process. This environment is different from usual KiPs in decision-making as open innovation is specifically of high strategic importance (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), particular risky (Banu et al., 2016; Herzog, 2008; Saura et al., 2022; Wu et al., 2013), and operates in highly dynamic environments (Banu et al., 2016). The proposed approach tries to investigate the proper functioning in a knowledge-intensive setting that heavily relies on (personal) experiences rather than historical data, and to which extent the study's result could be generalized to various contexts. Figure 3 shows the visualization of this connecting mix of elements that were used as an explorative case study to investigate the actual functioning.

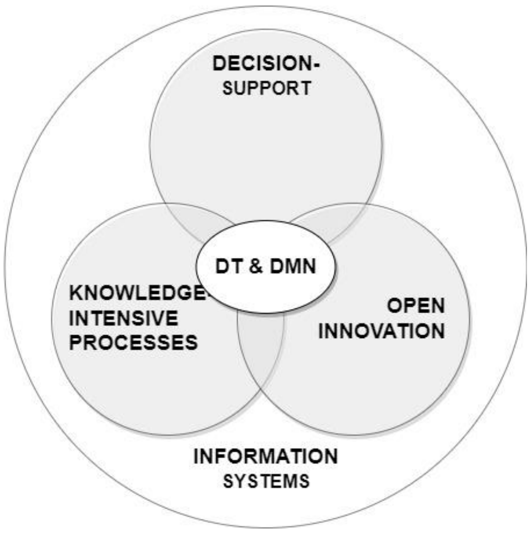


Figure 3. Recommended elements for Explorative Case Study

1.3 Report outline

This section explains the remaining outline of this report to clarify how the structure should be read for better understandability. The previous sections described the context of this research, but the research's core is described in the upcoming chapters. First, Chapter 2 gives the theoretical background, which is an overview of the relevant literature related to the research's subject(s). This theoretical background is based on a Structured Literature Review about *knowledge-intensive decision-making within open innovation.* The discussed topics are characteristics of KiPs, characteristics of open innovation, business process modelling, and machine learning. Next, Chapter 3 describes the research methodology, which is case study research methodology, so the case study design is central to this chapter. Moreover, this chapter focuses on how the approach can be implemented in the case using a combination of two frameworks for industrial settings. Thereafter, Chapter 4 shows the case study results, which effectively combine all sub-questions defined for the case at Shell. Additionally, the generalization is central to the last section of Chapter 4. Finally, Chapter 5 discusses both the case study conclusions and research conclusions. Moreover, this concluding chapter elaborates on the research's limitations and recommendations for both the business case and further academic research.

2 Theoretical background

This chapter provides additional background information about theories and methods that are applied in this study to increase the understandability of these theories and methods in the remaining parts of this report. A previously conducted Structured Literature Review provides a base for this theoretical background chapter and is created using a specific search strategy presented in Appendix A.

2.1 Introduction

This chapter aims to gain insight into the connection between knowledge intensity and open innovation to improve decision-making capability accordingly. The Business Process Management (BPM) field fulfils the need for processes with a focus on modelling and visualizing processes. On the other hand, BPM is not commonly used for data analytics, however, Machine Learning (ML) is. ML is part of the well-known concept of Artificial Intelligence, and is about using computer algorithms and is of key importance in accessing data (Ali et al., 2022). Therefore, the remainder of this chapter discusses the characteristics of KiPs, characteristics of open innovation, business process modelling, and ML techniques to give a complete overview.

2.2 Characteristics of knowledge-intensive processes

Over the past decades, business processes have been changing in point of view regarding success. They have been changing from success in terms of predictability and efficiency to success in terms of capability to adapt and treat unique customer requests, which require creativity and collaboration and rely on the knowledge of experts (Boissier et al., 2019). Typically, these processes are KiPs, which have several decision-making tasks and heavily rely on knowledge workers' expertise and experience (Venero et al., 2019). Knowledge workers are decision-makers with different backgrounds and can create and exploit specific domain knowledge to achieve business goals. Earlier research showed that the interpretation of KiPs barely differs, and the returning main character within these processes is knowledge. In the paper of Boissier et al. (2019), a successful KiP is adaptable for specific situations and treats unique customer requests instead of executing a predefined model. However, França et al. (2012) characterised a KiP as highly dependent on the involved knowledge in the minds, tasks, and activities of knowledge workers. These processes focus on knowledge conversion among knowledge workers involved in a business process execution (França et al., 2012). The definition stated in the paper of Isik et al. (2012) agrees with the beforementioned definitions and highlights that KiPs require both information and knowledge collection and use.

Di Ciccio et al. (2015) represented KiPs based on eight key characteristics namely Knowledge-driven, Collaboration-oriented, Unpredictable, Emergent, Goal-oriented, Event-driven, Constraint- and rule-driven, and Non-repeatable (Di Ciccio et al., 2015). First, KiPs are Knowledge-driven, which indicates that knowledge is the key driver and drives human decision-making and the process flow of actions and events. Di Diccio et al. (2015) continued by stating that process creation, management, and execution of KiPs occur collaboratively and, therefore, KiPs are Collaboration-oriented. Next, KiPs are Unpredictable as it might be unknown what the exact situation and context elements are. Therefore, the exact activity, event, and knowledge flow depend on unpredictable situations that may change or vary over time (Di Ciccio et al., 2015). As it is unknown beforehand in what way the actual process executes, the actual actions are determined stepwise and, therefore, KiPs are Emergent. Di Diccio et al. (2015) continued that KiPs are Goal-oriented as the process flow executes through defined goals or milestones that should be achieved. Moreover, KiPs are affected by multiple alternative events that influence decision-making by knowledge workers, which indicates that KiPs are Event-driven (Di

Ciccio et al., 2015). Next, Di Diccio et al. (2015) defined Constraint- and rule-driven as characteristic of KiPs as process participants may be influenced by constraints and rules that might drive decision-making. Finally, KiPs are Non-repeatable as the process execution has a unique nature that might significantly deviate from other process executions, which are hardly repeatable (Di Ciccio et al., 2015).

KiPs are upstream within organizations, however, their management might be difficult due to multiple challenges. Boissier et al. (2019) discussed reasons why the management of KiPs is associated with challenges in modern organizations. Besides the required focus on the characteristics of KiPs to ensure adequate management of these prosses, KiPs require attention to context-specificity, flexibility, and collaboration (Boissier et al., 2019). In addition, França et al. (2012) discussed challenges that largely correspond to the challenges mentioned by Bossier et al. (2019), but the authors focus specifically on subjective judgement by knowledge workers involved. KiPs typically comprise steps based on personal experiences, which might lead to undesirable consequences in organizations. Therefore, KiPs are complex and difficult to get structured and automated (França et al., 2012). On top of that, Isik et al. (2012) also defined challenges, however, the authors agree with the beforementioned challenges and conclude that KiPs require another modelling approach than non-KiPs.

2.2.1 Take-aways for study

This section briefly discusses the relevance of the abovementioned theoretical background regarding KiPs for this study. First, within business processes, the capability to adapt unique customer requests became more important and, therefore, business processes require creativity and collaboration (Boissier et al., 2019). These processes are KiPs and heavily rely on experts' knowledge implying both expertise and experience (Venero et al., 2019). Because these types of processes are different from non-KiPs, KiPs require another management and approach (Boissier et al., 2019; França et al., 2012). Therefore, the proposed approach of this study is carefully defined to suitably connect the characteristics of KiPs (Isik et al., 2012) as knowledge intensity is one of the core elements within this study.

2.3 Characteristics of open innovation

Nowadays, organizations are facing increasing complexity within their open innovation processes and are challenging decisions daily. Typically, knowledge is an unavoidable and fundamental resource for creating innovation and innovation is commonly seen as the combination of knowledge and ideas (Wu et al., 2013). Banu et al. (2016) stated that open innovation is a complex flow of knowledge, which implies a large number of actors in a highly dynamic environment. Wu et al. (2013) defined open innovation as a paradigm where internal and external ideas should be used by organizations as it improves an organization's performance. Banu et al. (2016) agree with Wu et al. (2013) and stated that open innovation aims to accelerate innovation and expand markets for external use of innovation (Banu et al., 2016).

Bagherzadeh et al. (2021), Herzog (2008), and Saura et al. (2022) defined properties that characterize open innovation. Table 1 shows these properties and both similarities and differences between the papers. The papers agree with each other's defined characteristics of open innovation, however, Saura et al. (2022) did not discuss the uncertainty of open innovation in their paper. As these papers have reasonable similarities in characterizing open innovation as KiP, the main characteristics of open innovation are accepted as uncertain, complex, of high strategic importance, and knowledge-intensive. Considering the latter, Saura et al. (2022) explicitly mentioned that social networks can be a relevant source of knowledge in organizations, which could also help improve open innovation.

Table 1. Properties of open innovation

Property \ Paper	Bagherzadeh et al. (2021)	Herzog (2008)	Saura et al. (2022)
Uncertain	X	X	-
Complex	X	X	X
Strategic importance	X	X	X
Knowledge-intensive	X	X	X

The capability to innovate is most certainly crucial to gaining a competitive advantage in today's competitive and fast-changing organizations (Wu et al., 2013). However, organizations should have the willingness and be open to accepting external knowledge before becoming innovative for which multiple challenges should be conquered. Recently, Saura et al. (2022) published a paper that defined multiple challenges to open innovation. First, open innovation might evoke negative attitudes towards external knowledge and cooperation through communication in organizations (Saura et al., 2022). Therefore, a manager's role and strategic support for open innovation are of huge importance. The greatest loss is that companies do not make sufficient effort to establish alternatives that manage open innovation as a possible key to success from both technological, organizational, and project decision-making perspectives (Saura et al., 2022). The attitude toward open innovation in organizations should be changed to the perspective that exchanging knowledge and cooperation could enhance creative innovation and boost knowledge generation (Saura et al., 2022).

Moreover, open innovation encounters risks and risk-taking encouragement is one of the fundamental elements of open innovation due to its innovative nature (Herzog, 2008; Wu et al., 2013). Additionally, Banu et al. (2016) defined eleven risks related to open innovation. The first two risks are the misalignment between innovation's objectives and an organization's strategic objectives, and an unrealistic innovation evaluation. Next, the third risk is the lack of human resource that is sufficiently skilled. The fourth, fifth and sixth risks are communication-related and are defined as follows: inefficient integration in the open innovation network, and ineffective communication within both the organization and with innovation partners. Next, innovation-related Key Performance Indicators are not sufficient used or developed in the evaluation process. The ninth and tenth risks defined by Banu et al. (2016) are poor management and the lack of markets as consumers' preferences may change. The last risk found by Banu et al. (2016) is the competitor's level of technology, which might be much greater than the organization's adequate capabilities. Additionally, Banu et al. (2016) highlighted the importance to model open innovation processes as it enables experts to identify contingency scenarios relevant to the process to uncover errors along with the process flow (Banu et al., 2016).

2.3.1 Take-aways for study

This section briefly discusses the relevance of the characteristics of open innovation for this study. First, organizations are facing increasing complexity within their open innovation processes where knowledge is unavoidable (Wu et al., 2013). However, open innovation processes are different from usual KiPs as open innovation is specifically of high strategic importance (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), and operates in highly dynamic environments (Banu et al., 2016). Additionally, open innovation faces multiple challenges and is particularly risky (Banu et al., 2016; Herzog, 2008; Saura et al., 2022; Wu et al., 2013). Therefore, the proposed approach considered the characteristics of open innovation to suitably connect with this environment. Moreover, the possibility of encountering the risks is captured by the approach to prevent misalignments, lacking understandability of stakeholders, insufficient skilled resources, and missing evaluation methods.

2.4 Business process modelling

Over the past decades, the complexity of business processes has been rapidly accelerating due to the increasing pace of the business world (Isik et al., 2012). Today's companies are extensively looking for solutions that structure and standardize their business processes in a well and easy way to monitor (Isik et al., 2012). Isik et al. (2012) discussed that at the same time as the increasing growth in organizations to differentiate themselves from their competitors, implying the increasing focus on expert knowledge within business processes, the attention to BPM tools and methodologies is increasing. Di Ciccio et al. (2015) defined the BPM field as an active area of research with high practical relevance. BPM is about the situation where several activities are necessary before the outcome is reachable such as a product or service provided to the market. The BPM field offers approaches to improve companies' workflow focusing on redesigning processes to achieve improved efficiency and effectiveness (Kluza et al., 2017).

2.4.1 Approaches

Achieving strategic and operational business goals is largely supported by the right decision-making along with the process execution (Bazhenova & Weske, 2016). BPM focuses on the design and control of business processes using techniques that require several sources of information (Di Ciccio et al., 2015). The BPM field offers three of the most commonly used process improvement standards offering support for process specifications founded by OMG (OMG, 2019). According to OMG (2019), these three commonly used standards within the BPM field are called the *triple crown* representing: Business Process Management and Notation (BPMN), Case Management Model and Notation (CMMN), and Decision Model and Notation (DMN). The remainder of this section analyses each approach in more detail and discusses its benefits and limitations accordingly.

2.4.1.1 BPMN

Earlier research showed that Business Process Model and Notation, BPMN in the remainder, is the most commonly used standard for designing business process models. It is used to easily design, manage, and realize business processes, which could be directly used by stakeholders (OMG, 2019). OMG (2019) defines BPMN diagrams as having sufficient precision to be able to get translated into software process components and are easily usable due to their independent notation in any environment (OMG, 2019). In addition, Kluza et al. (2017) described that within the widely used modelling notation BPMN, the process diagram is the most commonly used diagram as this type is sufficient for most cases. The process diagram describes the executed flow of operations to reach defined business goals (Kluza et al., 2017), which can be seen in Figure 4.

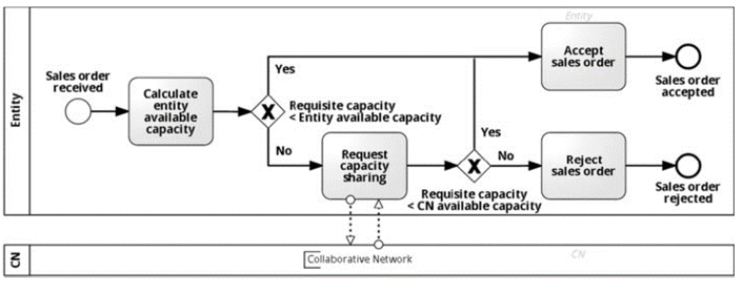


Figure 4. Example BPMN on 'Order Acceptance' (source: Biard et al., 2015)

2.4.1.2 CMMN

The first version of the Case Management Model and Notation, CMMN in the remainder, was launched in 2014 by the OMG group (De Carvalho et al., 2016). CMMN is an extension of BPMN and is both a meta-model and notation for modelling to express a case or multiple cases. However, CMMN has completely changed to a different modelling approach and can be better related to the Guard-Stage-Milestone (GSM) metamodel as CMMN relies on GSM constructs (Di Ciccio et al., 2015). Di Ciccio et al.

(2015) stated that CMMN has a strong connection to the business artefacts framework, which provides a methodology that is data-centric. Additionally, as described in De Carvalho et al. (2016), CMMN models are the standard for dealing with cases and a CMMN model has two existing phases. First, CMMN models have a design phase where segments in the case model are defined by the business analyst, so the case can be customized by the case management in runtime. Second, after the design phase, the process executes by the case manager along with the predefined tasks (De Carvalho et al., 2016). Figure 5 shows an example of a CMMN model on 'Emergency'.

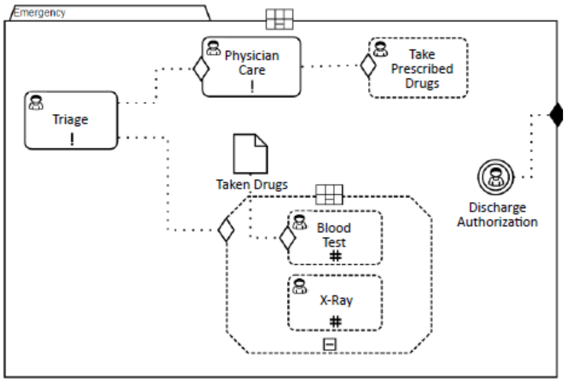


Figure 5. Example CMMN on 'Emergency'
(source: De Carvalho et al., 2016)

2.4.1.3 DMN

Decision Model and Notation, DMN in the remainder, is found to capture multiple gaps within BPM tools. First, the main goal of DMN is to provide a standard notation, which is understandable for all business users from business analysts to technical developers to business people. Within this notation, DMN is created to standardize a bridge for the gap between designing business decisions and implementing decisions and could be used besides the general BPMN. Next, DMN is created to enable the interchangeability of decision models over organizations via XML representations (OMG, 2019). According to Figl et al. (2018), DMN is about specifying both business rules and business decisions in a precise way. The authors continue by stating that DMN helps business users to control both their organization decisions and processes to reach more efficiency and effectiveness in terms of well-designed information and decision structures (Figl et al., 2018).

Figure 6 gives an example of a DMN model on 'Credit Sales' and visualizes three aspects of decisions: the decision requirements level, the decision logic level, and the expression language. First, the upper level consists of a decision requirements diagram (DRD), which presents relationships between decisions by information requirements. Second, the decision logic level is the middle part of the figure, which presents a single decision's logic by using a boxed expression. The most commonly used presentation for representing the decision logic level is a decision table, which defines production rules from input parameters to output parameters (Figl et al., 2018). Finally, the third and last element is the expression language, which is presented in the bottom part of Figure 6. The paper of Figl et al. (2018) stated that DMN standardizes Friendly Enough Expression Language (FEEL) as the expression language and the language S-FEEL can be used as a subset of expression language for use in decision tables. FEEL defines a syntax that can be used for expressions, which allows the description of decision logic by decision tables or other possible alternatives beforementioned (Figl et al., 2018). The bottom part of Figure 6 shows the decision logic of 'Credit Eligibility' using FEEL as expression language for expressing the rules in the decision table.

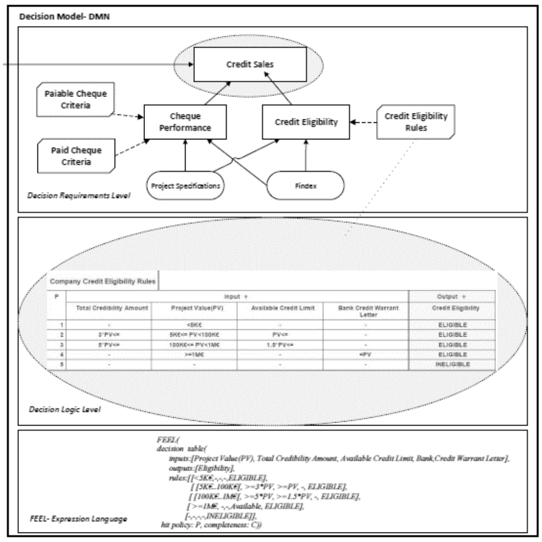


Figure 6. Example DMN on 'Credit Sales'
(source: Figl et al., 2018)

2.4.2 Benefits and limitations

Although all modelling approaches of the triple crown are suitable for process management support, there are differences between them regarding applicability within specific cases. BPMN is the standard notation for modelling business processes and is capable of describing end-to-end processes as a flow of tasks (Bazhenova & Weske, 2016). To perform this, Kluza et al. (2017) defined that BPMN 2.0 users can choose a process diagram, collaboration diagram, conversation diagram, and choreography diagram (Kluza et al., 2017). BPMN 2.0 defines more than 100 elements so process modellers can choose their own model's level of detail. The three basic levels, as many more extensions are possible to capture, of BPMN are a descriptive level, an analytical level, and an executable level (Kluza et al., 2017). However, BPMN does not support concepts as modelling rules so the underlying decision logic in BPMN is not (easily) interpretable from the process model, resulting in a decision logic that is harder to automatically derive (Bazhenova & Weske, 2016). Biard et al. (2015) agree with the abovementioned limitation of BPMN and argued that BPMN only enables investigation when a decision is made rather than how a decision is made (Biard et al., 2015). Therefore, DMN complements BPMN as DMN integrates this aspect within the possibilities of the modelling notation.

DMN fills in the gap in decision modelling and has multiple benefits. The first reason to use DMN is that DMN models improve speed and project cycles will become faster (Etinger et al., 2019). Next, DMN models increase the participation of business stakeholders and hidden relationships can be revealed by visualization through graphical logic (Etinger et al., 2019). Moreover, DMN models ensure flexibility, transparency, improved efficiency, and improved quality and compliance by standardizing processes and their decisions (Etinger et al., 2019). Besides the wide range of tools that DMN offers, the modelling notation is easily understandable. DMN fulfils the need for organizational decision-making and it is particularly useful when decisions are high-risky for the operations (Etinger et al., 2019). This is due to the possibility of expressing risk as category high, medium, or low by using one of the S-FEEL data types as expression language (Calvanese et al., 2018). The core value of DMN is modelling (human) decision-making, and both investigating requirements for automation and implementing this automated decision-making (Kluza et al., 2017). As described in Kluza et al. (2017) and Bazhenova & Weska (2016), DMN is often used in combination with BPMN, but can also be used alone standing. BPMN does not provide decision logic, but it provides rule tasks that can be linked to decision logic in other decision models created by DMN (Bazhenova & Weske, 2016). The papers of Biard et al. (2015) and Etinger et al. (2019) examine the complexity of DMN models and described a couple of drawbacks of the DMN modelling notation. First, uncertainty is one of the key elements that DMN is unable to deal with (Biard et al., 2015). Next, it is required to avoid large amounts of data by clearly understanding the business objective or business rules, which can be hard in some situations (Biard et al., 2015). Finally, Etinger et al. (2019) stated that the presentation in DMN's decision tables might be a bottleneck when using ML models, which will be discussed in the next section.

CMMN is a more data-related modelling notation and enables the modelling of different categories of flexible business processes, which provides tasks and constraints that should be taken into consideration along with the process execution (De Carvalho et al., 2016). The benefit of CMMN is that it defines potential upcoming times or so-called discretionary items. Once these discretionary tasks are defined by the business analyst during design time, the case manager is in turn allowed to change or plan a specific task triggering execution (De Carvalho et al., 2016). However, De Carvalho et al. (2016) stated, in addition, a couple of limitations of CMMN. First, within the CMMN modelling notation, an executed task cannot be re-executed. Moreover, CMMN does not provide an intrinsic definition of resources so the business analyst is not triggered to make a model that is simplified without resources definition, and neither other artefacts to deal with involved resources. The role of this emerging approach requires further (formal) investigation for evaluation in concrete situations (De Carvalho et al., 2016; Di Ciccio et al., 2015).

2.4.3 Take-aways for study

This section briefly discusses the reasoning behind the business process model used in this study accordingly. Based on all discussed approaches, benefits, and limitations, it could be concluded that the combination of BPMN and DMN is best suitable for the specific context. The combination of BPMN and DMN is the best modelling approach to represent knowledge-intensive decision-making within open innovation because this approach supports both knowledge-intensity and decision-making logic.

2.5 Machine learning

Today's organizations are facing increasing amounts of complex data available, which is identified in every business aspect and enables decision support by providing information and knowledge (Lu, 2018). This increasingly digital world affects businesses and jobs and provides highly complex datasets that are unable to analyse with traditional methods (Lu, 2018). This phenomenon refers to Big Data, which comprises three V's: (high) Volume, (high) Velocity, and (high) Variety. Lu (2018) described Volume as the amount of data created, Velocity as the speed of data-generating and transferring, and Variety as the number of variants in data and sources. Although the majority agrees with this basic expression of Big Data, Mittal & Sangwan (2019) defined an extended definition of Big Data. Any data can be considered Big Data when it is not only about Volume, Variety, and Velocity, but also about Veracity and Value, which comprises the 5V model (Mittal & Sangwan, 2019).

As the growth of data is increasing daily, the amount of challenges in data is growing as well, which has led to the increasing need for emerging technologies rather than traditional methods (Lu, 2018). As a result, organizations are motivated to use upstream technologies such as ML, and Czvetkó et al. (2022) stated that the number of ML projects is expected to double within a year. ML is part of the well-known concept of Artificial Intelligence, which improves businesses' data (mining) analytics to improve efficiency (Czvetkó et al., 2022). As described in Czvetkó et al. (2022), ML algorithms aim to recognize patterns in data to support decision-making, which makes decision-making in processes faster and more accurate. In more detail, ML is about predicting future outcomes on data, which is based on predefined rules, and no human intervention is involved. Therefore, the learning and refining of the algorithm are fully automatic (Czvetkó et al., 2022). This section aims to gain insight into ML techniques, where the first step is to understand the possible ML techniques, and the next step is to analyse both the benefits and the limitations of these techniques accordingly.

2.5.1 Techniques

ML techniques are distinguished based on two different segments, namely whether the ML algorithm is supervised or unsupervised and whether the output of the ML algorithm is continuous or categorical (Lu, 2018). This distinction results in four different ML categories: Classification, Clustering, Regression, and Dimension Reduction. Figure 7 shows the two different segments and the four different categories accordingly. The most crucial segment within ML is whether the ML algorithm is supervised or unsupervised, so this distinction is central in the remainder of this section.

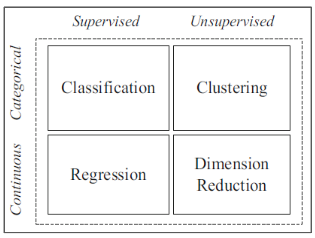


Figure 7. Categories of ML algorithms (source: Lu, 2018)

2.5.1.1 Supervised learning

Supervised learning algorithms versus unsupervised learning algorithms are the most crucial distinction within ML, and this section focuses on this first type of ML algorithm. Supervised learning algorithms differ from unsupervised learning algorithms by mapping input parameters to their associated output supported by numerical parameters (Mittal & Sangwan, 2019). The well-known categories within this segment are Classification, mostly used for categorical or discrete output, and Regression, mostly used for continuous output. Within these two categories, many alternatives are possible to use as ML algorithms, however, this SLR focuses on only a few of them due to clarity and limited scope. Whereas Decision Trees (DT), Support Vector Machines (SVM), and K-Nearest Neighbor (k-NN) are examples of Classification models, the well-known example of a Regression model is Linear Regression (Ali et al., 2022; Mittal & Sangwan, 2019). Within supervised learning algorithms, most methods are multi-use methods as these methods are suitable for both Classification and Regression situations, e.g. DT, SVM, and k-NN fulfil both Classification and Regression problems.

After introducing the general background of supervised learning methods, some of these methods will be briefly explained in more detail. First, SVM is, as earlier described, a multi-use method where the Classification variant is the so-called Support Vector Classification (SVC) and the Regression variant is the so-called Support Vector Regression (SVR). SVC analyses data for classification by handling learning models with learning algorithms, which are designed to categorize multiple data types from multiple disciplines (Ali et al., 2022). The SVM algorithm fulfils best in cases where classes are divisible within the dataset, and it is particularly suitable for significant binary cases (Ali et al., 2022). Next, k-NN is part of the most fundamental programming algorithms and makes ML easier to implement (Ali et al., 2022). Ali et al. (2022) described that the idea of k-NN is that input values are used to simulate output values, and k-NN based the categorized information on its neighbor's ranking by classifying new data points as the similarity of previous data points as a base. The 'k' in k-NN stands for the number of nearest neighbors that are used for classifying new data points (Ali et al., 2022).

Moreover, DT is a more graphical model based on the branching method that illustrates possible decision outcomes (Etinger et al., 2019). The papers of (Etinger et al., 2019) and (Lu, 2018) described the DT model. A DT model splits data into smaller sub-categories until the sub-categories are uniquely determined (Etinger et al., 2019). As the DT is a graphical model, it is based on multiple standard features, and Figure 8 shows an example of a DT. The seven rectangular boxes are the leaf nodes, representing the outcomes – or classes, the internal nodes stand for a test on the variable, and the outcomes for the test are represented by each brand (Lu, 2018).

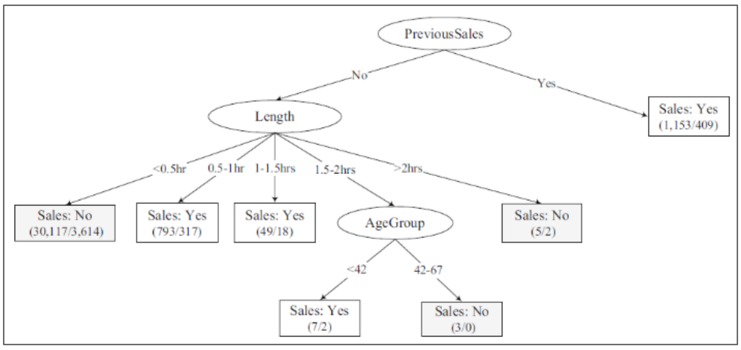


Figure 8. Example DT model (source: Lu, 2018)

2.5.1.2 Unsupervised learning

In addition to the previous section, this section focuses on the other type of ML algorithms, namely unsupervised learning algorithms. Unsupervised learning algorithms differ from supervised learning algorithms by learning patterns from a given dataset (Mittal & Sangwan, 2019). The well-known categories within this segment are Clustering, mostly used for categorical or discrete output, and Dimension Reduction, mostly used for continuous output. Again, within these two categories, many alternatives are possible to use as ML algorithms, however, this SLR focuses on only a few of them due to clarity and limited scope. Whereas examples of Dimension Reduction are quite implicit in current literature, the most commonly used examples of Clustering algorithms are k-means, DBSCAN, and OPTICS algorithms (Mittal & Sangwan, 2019). Clustering algorithms are based on data mining analysis techniques that are used for grouping data instances based on similarities in metrics (Lu, 2018). Earlier research showed that the most commonly used unsupervised learning technique is k-means, therefore, this technique will be briefly discussed in more detail.

The paper of Zineb et al. (2021) is recently published and the authors described the k-means algorithm clearly. For the k-means algorithm, knowledge of the number of groups is required in advance and the elbow method can be applied to determine this (Zineb et al., 2021). The elbow rule is based on a figure where the X-axis stands for ‘K’, which represents the number of the cluster and the Y-axis stands for the sum of squares for the clusters. The best suitable cluster is chosen at the moment when the highest reduction in the sum of squares takes place, i.e. extreme decrease (Zineb et al., 2021). Therefore, the elbow rule can be interpreted as a method where the bend (just like an elbow) in the figure indicates the solution.

2.5.2 Benefits and limitations

Although all ML techniques are perfectly suitable for recognizing patterns in data to support decision-making, there are differences between them regarding applicability within specific cases. After explaining the possible techniques for supervised learning algorithms and possible techniques for unsupervised learning algorithms, this section dives into deep by analysing the benefits and limitations of the techniques accordingly. First of all, the general benefit of ML is that it obligates data analysts to structure and analyse their datasets appropriately. This is because ML algorithms do not work properly once datasets are not well prepared, which lacks extracting deep insights that are possibly hidden in data. Using raw data to perform analyses most certainly leads to incorrect decision-making due to poor data quality (Mylavarapu et al., 2019). Data quality is the key factor that influences data analyses in multiple ways, but consistency is one of the most important dimensions (Mylavarapu et al., 2019). However, the major drawback of this part is that ML algorithms are trained in such a way on a certain dataset that the algorithm might be useless and unsuitable for another dataset (Mittal & Sangwan, 2019).

In the paper of Etinger et al. (2019), a method for creating DMN decision tables from a DT model is presented, which is automatically generated by a DT Classifier. Earlier research showed that Python enables the module scikit-learn as a well-fitting ML library for building DT. Every rule (row) in the decision table represents a DT’s leaf (Etinger et al., 2019). The benefit of this DT is that this ML model is easily interpretable, however, the drawback is that DT might lack accuracies compared to black-box models (Sachan et al., 2020). The paper of Mittal & Sangwan (2019) stated that DT is not appropriate for big data analytics and that SVM is suitable for databases of reasonable size. k-NN can only be used when data is normalized, noise-free, and consequently labelled, which makes it more difficult to apply in certain scenarios and finding a neighbor is not easy as small values of 'k' might highly influence the result (Ali et al., 2022). Moreover, ML algorithms are relatively easy to evaluate, and an example is presented in the papers of Sachan et al. (2020) and Zineb et al. (2021). A classification report is one of the possibilities for evaluating ML algorithms. This matrix presents four commonly used metrics: accuracy, precision, recall, and f-score. Zineb et al. (2021) compared multiple ML models on these four metrics and found that SVM and DT have the same scores on the four metrics for their specific research case.

2.5.3 Take-aways for study

This section briefly discusses the reasoning behind the ML technique used in this study accordingly. Based on all discussed techniques, benefits, and limitations, it could be concluded that DT as a supervised ML model is perfectly suitable for analysing decision-making data because a DT model is highly interpretable and understandable. Moreover, DT is linked to the previously chosen combination of BPMN and DMN as a business process model because it supports decision tables in the DMN model(s), which could be fully automated. The next section concludes the theoretical background by discussing the overall take-aways of all previously described sections together.

2.6 Conclusion

In summary, this chapter described the characteristics of KiPs and open innovation. Although some characteristics of open innovation equal characteristics of KiPs such as its uncertainty, complexity, and knowledge intensity (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), open innovation differs from usual KiPs. Open innovation is specifically of high strategic importance (Bagherzadeh et al., 2021; Herzog, 2008; Saura et al., 2022), particular risky (Banu et al., 2016; Herzog, 2008; Saura et al., 2022; Wu et al., 2013), and operates in highly dynamic environments (Banu et al., 2016). Moreover, this chapter investigated current knowledge about business process modelling. It is found that the combination of BPMN and DMN is the best modelling approach to visualize knowledge-intensive decision-making within open innovation. This supports both knowledge-intensity and decision-making logic. DMN enables the possibility to extract decision models from event logs. DMN fulfils the need for organizational decision-making and it is particularly useful when decisions are high-risky for the operations (Etinger et al., 2019). This is due to the possibility of expressing risk as category high, medium, or low by using one of the S-FEEL data types as expression language (Calvanese et al., 2018). Hence, DMN is perfectly suitable for open innovation processes as they are risky. However, this is just a presentation tool, which indicates that data analysis concerning open innovation is still lacking. To fill in this gap, ML techniques are identified based on current literature, and it is found that DT is perfectly suitable for analysing decision-making data. DT is highly interpretable and understandable and linked to the combination of BPMN and DMN as it supports decision tables in DMN model(s).

The combination of BPMN and DMN with a Decision Tree (DT) is not unknown in scientific knowledge. This research continues on earlier work by building on existing approaches and evaluating these approaches in a business case specific to the open innovation environment. The expected contribution is combining the DT model to derive decision logic at decision points via decision-mining combined with the quite new DMN modelling approach as a visualization tool for the mined decision logic in a knowledge-intensive open innovation process. As earlier mentioned, this environment is different from usual KiPs. The proposed approach tries to investigate the proper functioning in a knowledge-intensive setting that heavily relies on (personal) experiences rather than historical data, and to which extent the study's result could be generalized to various contexts.

3 Methodology

This section describes the study's methodology, which is based on case study research methodology as the study is an explorative case study. Therefore, this chapter discusses the case study design to execute the research and is structured based on the case study research design discussed by Yin (2009). Figure 9 shows the methodology used in this study, which is numbered from steps 1-8 to make it easier to refer to those execution steps in the remainder of this report. The mix of elements of the case study is shown in the middle of the figure as the core element. The case study aims to identify the first steps toward a general approach that improves decision-making capability (the ability for supporting decisions based on historical insights) within knowledge-intensive innovation processes through the case at Shell. Hence, steps 2-7 are used to answer the main research question (step 1) and to investigate to which extent the case study could be generalized (step 8) to other contexts accordingly.

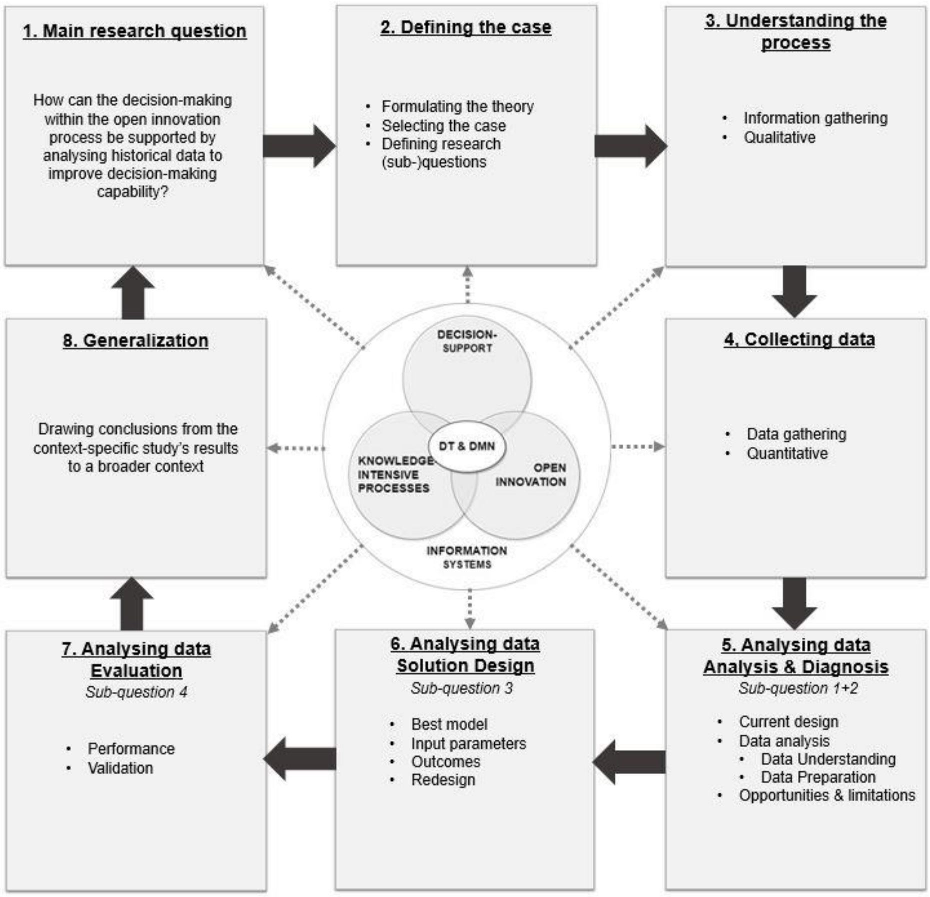


Figure 9. Case study design

3.1 Defining the case

First, the definition of the case is the preliminary stage of the case study research methodology, which is about formulating the theory, selecting the case, and defining research questions (Yin, 2009). These elements are already discussed in Chapter 1. One of the core elements of the case definition is the problem statement, which is already discussed in section 1.1.1 and supported by other previous sections. The problem statement is supported by sub-questions that should be addressed along with the research to reach a solution. The upcoming sections of the methodology describe the strategy to address the sub-questions to finally answer the main research question.

3.2 Understanding the process

As part of the fieldwork and analysis stage (Yin, 2009), this section describes the goal and the structure of information gathering. This section discusses the understanding of the process for the case, which refers to the qualitative part of the data collection method. This is a preliminary step to properly collect data accordingly, which is described in the next section. First, information is gathered via interviews with experts. Thereafter, this information is used to present the current process via process mining using the selected sources. The remainder of this section discusses both elements in more detail.

First, a combination of structured and unstructured interviews was used because this gained the best information possible to develop the process model due to multiple reasons. First, the interviewees could describe the process of the current situation, including opportunities and limitations within the process, in their way without knowing what the crucial parts for the expert were. Second, critical questions were pre-defined to gain information that is crucial for the research to prevent a lack of value-adding data. Furthermore, the interviews were transcribed in such a way that missing information or perspectives are discussed in another interview with the same interviewee. The interviewees involved did not have complete overviews of the current situation and their knowledge is based on (personal) experiences. Therefore, multiple disciplines were involved among the interviewees to prevent biased information, but the most important discipline is the GC department itself. Due to confidentiality, the roles of the experts involved cannot be explicitly stated, but the four experts involved in this research represent four different high-level functions. Almost all experts got their PhD in their expertise field (mostly electrical/chemical engineering) with significant professional experiences. All experts involved are stakeholders of Shell's open innovation process but with different backgrounds to approach the helicopter view rather than the centralized perspective.

Communication took place with one or two persons at a time to ensure focus via e-mail and MS Teams because barely any meeting was face-to-face due to COVID-19 and internationality. The combination of open- and pre-defined questions resulted in an unbiased but all-encompassing contribution to the understanding of the process. This gained information resulted in an understanding of the current situation regarding the defined scope described in section 1.1.3. Based on the information gained via the interviews, the current design was created and visualized via DMN (version 1.2) in Signavio. This current design formed a base to investigate and collect evidence for the limitations within the process. Interviews were the core source of information gathering, but the evidence for the issues is double-checked. This is done using self-directed analyses in the database to check the information collected, which is described in more detail in the next section.

3.3 Collecting data

As another part of the fieldwork and analysis stage (Yin, 2009), this section describes the goal and the structure of the data gathering. This section discusses the collection of data for the case, which refers to the quantitative part of the data collection method. After understanding the process, this part of the research methodology is to collect the (potential of the) available data in Shell's open innovation process. As a result, both the relevant available data for the process and the opportunities and limitations within the process of the current situation were investigated. Hence, possible areas for improvement in the current design were identified. This was done by evaluating projects in the GC's database by tracking project progress, recognizing patterns (exploring good and less good ones), and identifying problems.

Next, the case study focuses on discovering decision-making processes for open innovation projects where ML was used to reveal decision logic at decision points. The ML model was made using Python and the input for the ML model was a well-prepared dataset in MS Excel (converted to a readable CSV document by Python). Specifically, a Decision Tree (DT) Classifier, using the Python module Scikit-learn, was chosen to use as an ML model. The reason for this choice is that a DT perfectly suits analysing decision-making data, which is in more detail discussed in section 2.5 compared to other ML options. DT is highly interpretable and understandable and linked to the combination of BPMN and DMN as it supports decision tables in DMN model(s). Moreover, a DT perfectly fits the dataset as labelled data is available. The data is labelled since the output, whether a project was successful or not successful, is accessible and visible.

The DT relies on some input values which are related to characteristics in the GC's database. Therefore, the characteristics were carefully chosen as they should add value to the business use, which means that Shell's preferences are the guiding principle. Moreover, characteristics were also driven by data because a characteristic that has a one-sided value of 99 per cent does not make sense. This example of inconsistency was captured by a variance threshold and section 4.1.3 explains this type of data analysis in more detail. As a result, the data available were thoroughly analysed and resulted in extensive data analyses that gave insights into major data challenges.

3.4 Analysing data

To analyse the data and develop an approach for the case accordingly, a combination of two commonly used frameworks for problem-solving in industrial settings was used. These frameworks are well-known frameworks used to design improvement projects and help to analyse these projects in a problem-solving manner. The key methodology to implement the approach in the case is structured according to the Van Aken framework, which has a process-related perspective. Additionally, this study covers a data-related perspective, and this part is covered by the CRISP-DM framework. First, the framework of Van Aken is widely used as a problem-solving methodology in organizations to design practical business improvement projects based on various aspects and steps (Van Aken et al., 2012). Van Aken et al. (2012) defined a framework that considers five problem-solving phases: Problem Definition, Analysis & Diagnosis, Solution Design, Implementation, and Evaluation. Next, the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework is a more data-related framework, which guides data mining efforts in an industry-proven manner (IBM, 2021). The CRISP-DM framework consists of six phases, which are interrelated to each other, but this study only took into consideration the Data Understanding phase and the Data Preparation phase. Figure 10 shows the abovementioned methodology consisting of a combination of these two frameworks.

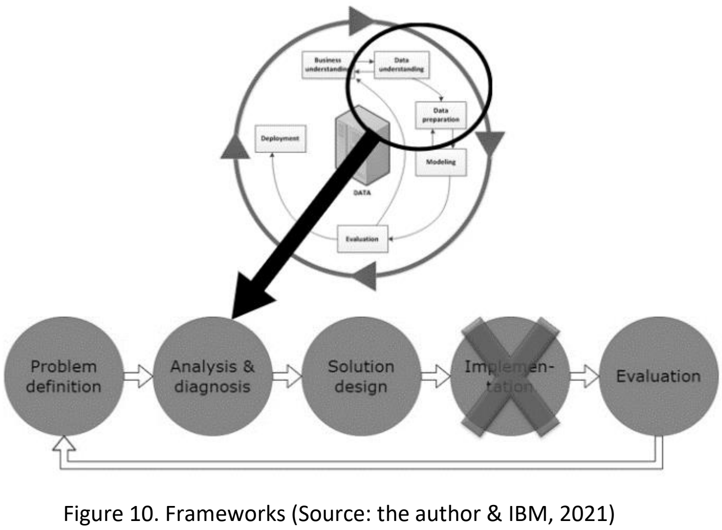


Figure 10. Frameworks (Source: the author & IBM, 2021)

Only three phases were used to implement the approach in the industrial case because two phases are irrelevant. First, the Problem Definition phase is irrelevant as this overlaps the case definition as the preliminary stage of the case study research methodology. Next, the Implementation phase is also irrelevant as there was no Implementation phase of the Van Aken framework due to the timeframe and the dependency on larger-scale research. The end of this case study is a validation of the new model as the redesigned process but no full implementation. Hence, the three phases that served as a base to develop and implement the approach in the case are Analysis & Diagnosis (supported by CRISP-DM's phases Data Understanding and Data Preparation), Solution Design and Evaluation. The upcoming sub-sections discuss all phases separately.

3.4.1 Analysis & Diagnosis

Analysis & Diagnosis is the second phase of the Van Aken framework, where the goal is to better specify the nature of the opportunity to increase the possibility to grasp this opportunity successfully (Van Aken et al., 2012). This phase is related to sub-question one and two: '*How is Shell's open innovation process currently designed?*' and '*What are the opportunities and limitations in Shell's open innovation process?*'. This phase of Van Aken is reinforced with two phases of the CRISP-DM framework but has separate purposes. The Van Aken framework is the framework at the higher level for the general understanding of the project and the CRISP-DM framework is the baseline for the data-related parts within this combination of frameworks. The CRISP-DM framework was used as guidance for the data-related parts of the project. As earlier explained, this study focuses on the Data Understanding phase and the Data Preparation phase of the CRISP-DM framework. According to IBM (2021), the Data Understanding phase is about getting basic insights into the data by activities such as collecting data and quality analysis. The Data Preparation phase is the next step where raw data changes into a data set that is ready to use because this step filters on inconsistencies such as outliers (IBM, 2021). The earlier discussed sections 3.2 and 3.3 formed a base for this Analysis & Diagnosis phase of analysing the data. Based on the abovementioned indications, the Analysis & Diagnosis phase of analysing data results in the following elements in the case study results chapter (Chapter 4): a current design, data analysis, and opportunities and limitations within this current design.

3.4.2 Solution Design

The next phase of the Van Aken framework is the Solution Design phase and is related to sub-question three: '*How can decision-making capability in Shell's open innovation process be improved?*'. After gaining knowledge about the current situation of the process and the opportunities and limitations in this process, this phase proposed a redesign as a possible solution to the problem statement. The main goal of this sub-question is to create a redesign that should improve decision-making capability. The current design could be interpreted as the design where the problem still holds, however, the redesign offered a solution approach for the problem stated.

Two different models (decision model and expert knowledge model) were combined as the proposed solution design. Within this combination of models, the decision model serves as a filter model for the expert knowledge model. Using the decision model (based on an ML's DT), projects could be rejected in an earlier stage before going into the expert knowledge model, which requires knowledge-intensive decision-making from Shell's experts. The decision model is based on historical data and the expert model is on the expert knowledge that is required for the process because of its knowledge intensity. The decision model was created using MS Excel (and converted to CSV) and Python, and the expert model was created using Signavio (DMN 1.2). Especially, the DMN DRD was created in Signavio's decision modeller using the basic DMN palette, the decision table was created in

Signavio's decision manager, and adding the DMN model to the BPMN model was done in Signavio's process editor. The earlier discussed sections formed a base for this Solution Design phase of analysing the data. Based on the abovementioned indications, the Solution Design phase of analysing data results in the following elements in the case study results chapter (Chapter 4): the model(s), input parameters, outcomes, and the process redesign.

3.4.3 Evaluation

Evaluation is the last phase of the Van Aken framework, where the goal is to evaluate the model's results compared to the business value (Van Aken et al., 2012). This part focused on the value of the models for meeting the business goals defined at the preliminary stage of the case study research methodology and is related to sub-question four: '*How can the proposed solution design be evaluated?*'. It is crucial to investigate whether the redesign is an improvement compared to the current design. This is a difficult part as insights in (performances of) decision-making capability were not known at Shell. Expert knowledge was the only option to use where Shell's experts were the core value.

The method for evaluating the proposed solution design used a Likert scale, which ranks five different questions from poor to excellent. All experts involved are the stakeholders mentioned in section 3.2 and the format used for the project evaluation can be seen in Appendix C. The questions allowed the experts involved to thoroughly understand the project. The topics of the five evaluation questions were: well-organized models, correspondence to practice, hands-on approach, old versus new model, and project's results versus business value. Additionally, experts could explain the ranking or raise awareness using optional fields to clarify their opinions. This expert knowledge was in addition to the model's results to validate that the model is an improvement. The study's results should support Shell's experts with their decision-making, especially in improving decision-making capability. The evaluation aimed to investigate to which extent the study's results are meaningful to Shell to be able to conclude if the model is satisfactory for business use in the specific case. To reach this, the impact of the identified changes was analysed in the redesign using the earlier mentioned evaluation method and steps for further research were suggested in the recommendations section accordingly. Here, the last element of the fieldwork and analysis stage (Yin, 2009) is described, which formed a base for the conclusion stage discussed in Chapter 5. Based on the abovementioned indications, the Evaluation phase of analysing data results in the following elements in the case study results chapter (Chapter 4): performance and validation.

4 Results

This chapter describes the results of the explorative case study conducted at Shell. The structure of this chapter follows the previously described combination of frameworks (Van Aken and CRISP-DM) to develop an approach for the case accordingly. All phases will be discussed separately consisting of the required elements per phase referring to the steps in Figure 9 accompanied by introductions and interim conclusions. To finalize this chapter, the fourth section discusses to which extent the case study could be generalized.

4.1 Analysis & Diagnosis

This section refers to step 5 of Figure 9 and comprises the Analysis & Diagnosis phase of the Van Aken framework, where the goal is to increase the possibility of successfully grasping the opportunity by better specifying its nature (Van Aken et al., 2012). This phase is related to sub-question one and two: '*How is Shell's open innovation process currently designed?*' and '*What are the opportunities and limitations in Shell's open innovation process?*'. This section uses multiple sections to give a complete overview of the Analysis & Diagnosis phase and consists of an introduction, current design, data analysis consisting of CRISP-DM's phases Data Understanding and Data Preparation, opportunities and limitations, and a conclusion.

4.1.1 Introduction

This section aims to form a starting point for developing a redesign as a possible solution to the problem statement. Therefore, the goal of the Analysis & Diagnosis phase is to gain insights into the current situation of the open innovation process at Shell and the opportunities and limitations within this process. As explained in the methodology chapter, this phase of the Van Aken framework is reinforced with two phases of the CRISP-DM framework, which provides a baseline for the data-related parts within this combination of research frameworks.

4.1.2 Current design

This section shows how the open innovation process at Shell currently is designed based on the methodology described in Chapter 3. As introduced, the current design uses the BPMN and DMN modelling approach to properly visualize the current design. The current process flow is depicted in Figure 11, but the labels have no description nor indicated description due to confidentiality reasons. The explanations of the objects used in both the BPMN model and the DMN model can be seen in Appendix B.

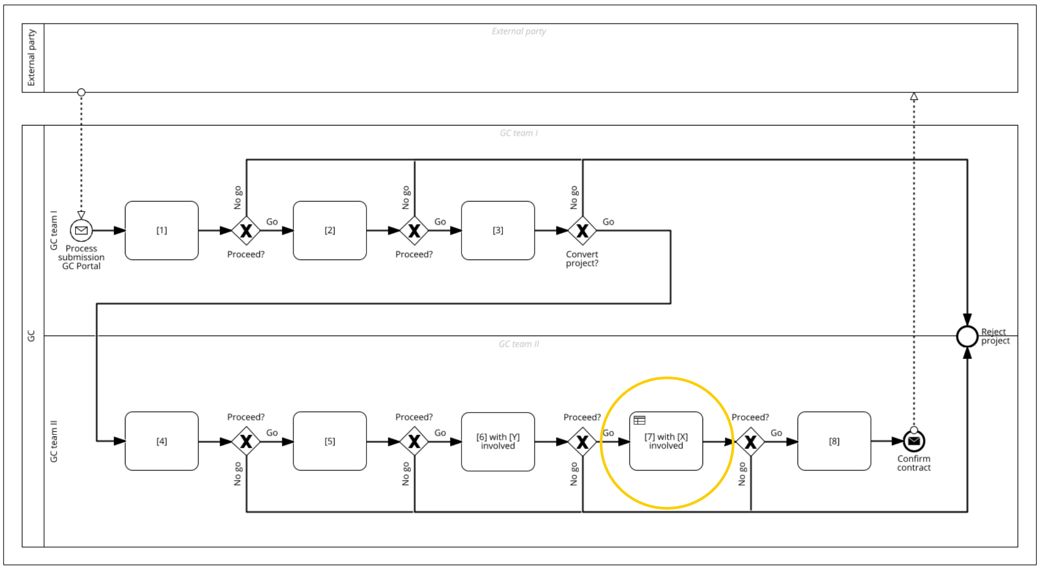


Figure 11. Current design via BPMN

First, the process is divided into two lanes of the GC process, one lane represents GC team I as actors and the other lane represents GC team II as actors. The composition of the teams differs, but how they differ cannot be explicitly stated due to confidentiality. The core of the difference is that team I handles the [M] phase whereas team II handles the [N] phase, and in both phases is another composition of roles involved required. The process starts when a submission from an applying external party, via the online submission form, enters the GC Portal. This type of project is called a [M], which is an incoming technology development that is still in the intake phase. Due to confidentiality, the stages cannot be explicitly discussed nor an indicated description of the stages. All [M]'s proceed with the next stage or are followed by a reject, which is both fully manual done by Shell's experts and not automatically. A rejected [M] is followed by actions by the GC team but this is out of scope. The [M] phase ends because the [M] is either rejected or started with the [N] phase. To clarify, a [N] is a technology development that successfully passes the intake process and proceeds with more detailed analyses. Ultimately, [X] in stage [7] is the most critical stage where expert knowledge is highly important that decides an important approval for a project to proceed with stage [8] which is the end of the scoped process. Therefore, this current activity, yellow circled in Figure 11, is chosen as the decisive moment that shows the fundamental part of expert knowledge involved. Both [X] and [Y] represent assessments by a group of experts, however, [Y] is the preparation of the assessment and [X] is the actual assessment.

This expert knowledge model is for the first time modelled and visualized and Figure 12 shows this using DMN DRD. The DMN DRD shows underlying decision specifications for stage [7] in the BPMN model. All the decision specifications presented relate to the BPMN stage [7]. This model is the expert knowledge model in the rest of the study, which requires knowledge from Shell's experts that are required for the process because of its knowledge intensity. Although DMN has more levels besides the DRD level as discussed in Chapter 2, the decision logic level of this model is out of scope. This part relates to one of the major limitations of the study and is discussed in more detail in section 5.3. Nevertheless, this expert knowledge model is created and assigned to clearly show for what expert knowledge model the proposed solution decision model serves as a filter. The DMN DRD model shows what informs the decisions in and for [X], which is again the most crucial expert knowledge model in the process. The determination of approval by [X] is supported by both the support of [X] and the approval of [Y]. It is important to note that the approval by [Y] happens sequentially and not parallel. Hence, once [Y] did not give their approval for the particular open innovation project, there is no determination by [X] happening. As shown in the BPMN process flow, once a project is rejected after [Y], it never reaches [X]. However, when [Y] does give their approval for a particular open innovation project, this approval is input for the determination of the approval by [X].

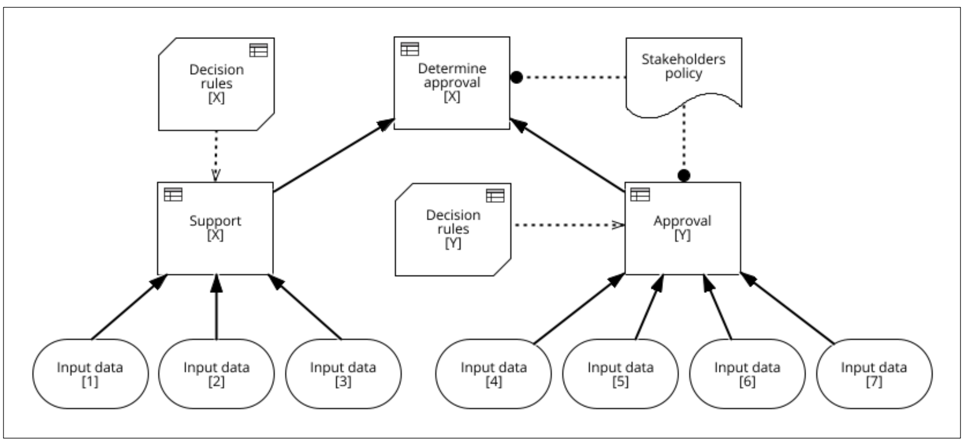


Figure 12. Expert knowledge model via DMN DRD

Two different decision rules are mentioned in the DMN DRD, but the distinction between [Y] and [X] is important because they differ significantly in people involved, documents, importance, and guidance. Next, the input data elements are different for support [X] and approval [Y] as support [X] is supported by input data elements [1], [2], and [3]. Next, approval [Y] is supported by input data elements [4], [5], [6], and [7]. Based on all these elements, the determination of the approval by [X] happens, which results in the execution of the BPMN stage [7]. This expert knowledge model is a crucial part of the process as it is the last step before proceeding with stage [8]. The resources of [X] are limited, which implies that Shell's experts have limited time, budget, and working hours (implying people resources). [X] deals with all incoming open innovation projects that reached that stage, i.e. all open innovation projects that are in the [N] phase and are not rejected along with the process flow as shown in the BPMN model. However, no existing (decision) model considers insights that support Shell's experts by their decisions, particularly not using historical data.

Currently, the decision-making within Shell's open innovation process fully relies on experts' knowledge that is based on an expert's past experience rather than using historical data in GC's database. As of 2020, the GC database had a major update concerning data quality because the fields in the online submission form have been made mandatory. Therefore, applying external parties had the obligation to fill in the fields of the online submission form rather than left some fields empty which results in missing values in the GC system. As a result, an analysis is done to check whether this update results in significant findings in differences in data quality, which will be discussed in the next section. Based on this analysis, it was found that the initial dataset collected consists of two different datasets. The first dataset, or the so-called dataset *before GC update 2020*, consists of open innovation projects from 2015, 2016, 2017, 2017, and 2019. The second dataset, or the so-called dataset *after GC update 2020*, consists of open innovation projects from 2020, 2021, and 2022.

The current decision-making process leads to the decision outcomes that are separately presented in Figures 13 and 14 based on the abovementioned distinction. Whereas Figure 13 shows the decision outcomes based on the dataset *before GC update 2020*, Figure 14 presents the decision outcomes based on the dataset *after GC update 2020*. The converted projects, implying successful projects, are significantly higher in the dataset *before GC update 2020* compared to the dataset *after GC update 2020*. The reason for this should be further analysed in the future when more data points are available in the *after GC update 2020* dataset as this dataset only consists of 2,5 years instead of 5 years in the *before GC update 2020* dataset. As earlier mentioned, the decision-making resulting in these decision outcomes fully relies on experts' past experiences. Therefore, the decision-making requires better pillars to build on future decisions that consider historical data to extract potential failure or success criteria.

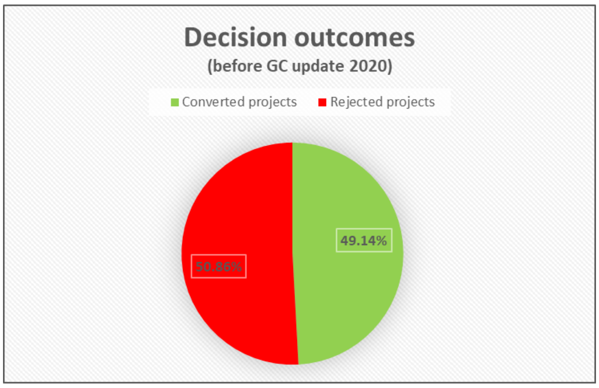


Figure 13. Decision outcomes *before GC update 2020*

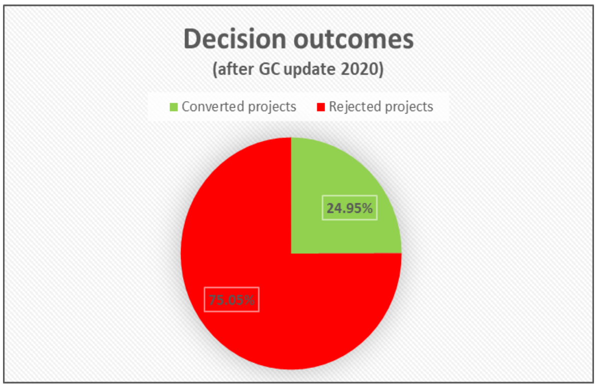


Figure 14. Decision outcomes *after GC update 2020*

4.1.3 Data analysis

The CRISP-DM framework is explicitly involved in data analysis concerning current activities by analysing opportunities and limitations of historical data within the process. The two phases involved, Data Understanding and Data Preparation, are separately discussed in the upcoming sub-sections.

4.1.3.1 Data Understanding

The first CRISP-DM data analysis phase involved was the Data Understanding phase. This phase was about collecting and verifying the data quality to get basic insights into data (Czvetkó et al., 2022). The output of the Data Understanding phase was the collection of initial data that is described, explored, and verified (Czvetkó et al., 2022). The Data Understanding phase was divided into multiple sub-sections to provide a complete understanding of the data. The sub-sections are data collection and data description, which are separately discussed below.

4.1.3.1.1 Data collection

This section explains how the study's data was collected and which elements shape the dataset. First, the data was collected via the GC's database, which deals with all data of Shell's open innovation process and contains detailed information about the characteristics of projects and the evaluations of decision-making over time. As of 2017, the data in GC's database is structured and could be of large queries depending on the filter area. The projects of this database are Shell-wide, which ensures that many people are working together, many alternatives are possible, and a high(er) budget is available compared to specific department projects. Shell's experts use this database to collect, identify, analyse, evaluate, and store information concerning all incoming innovation projects. The database contains available information from previous years, which indicates a well-organized collection of historical data. To ensure the selection of relevant data, the data collected fits the problem definition and scope described in Chapter 1. Therefore, open innovation projects from other innovation processes than GC were filtered out and the same holds for open innovation projects coming into the system other than the online submission form for external parties. Based on available data and according to the problem definition and scope, Table 2 presents the data attributes and class label (which is the output of the Decision Tree). Table 2 shows that nine relevant attributes are available and selected for this study plus one feature corresponding to the class label. Many fields of the table are censored due to confidentiality to prevent subtracting the attributes based on the attribute's characteristics.

Table 2. Overview of data collected

Nr.	Attribute	Type of variable		Possible values	Description
		Nature	Variant		
1	[A]	*Confidential*	*Confidential*	*Confidential*	*Confidential*
2	[B]	*Confidential*	*Confidential*	*Confidential*	*Confidential*
3	[C]	*Confidential*	*Confidential*	*Confidential*	*Confidential*
4	[D]	Categorical	Binary	Yes, No	*Confidential*
5	[E]	Categorical	Binary	Yes, No	*Confidential*
6	[F]	Categorical	Binary	Yes, No	*Confidential*
7	[G]	Categorical	Binary	Yes, No	*Confidential*
8	[H]	Categorical	Binary	Yes, No	*Confidential*
9	[I]	Categorical	Binary	Yes, No	*Confidential*
10	Status (= class label)	Categorical	Binary	Yes, No	Output: converted or rejected

4.1.3.1.2 Data description

Based on the previous introduction about the data collected, this section explores the data in more detail and describes the quality of the data. The upcoming first part describes the data mostly using the dataset size and thereafter, the quality of the data(set) is described. The dataset, based on the abovementioned nine attributes plus one class label, consists of 10 columns. However, the number of rows significantly deviates between the attributes and some attributes have only half the records as others. The reason for this is the tremendous differences in populated records of the attributes. Therefore, the maximum dataset shape is 10 columns by 1151 and the minimum dataset shape, i.e. taking only the populated records for all attributes, is depending on the combination of populated records for the 10 columns. Table 3 shows the abovementioned characteristics of the dataset shape.

Table 3. Overview of dataset shape

Dataset columns: attributes	**Dataset rows: Populated records**
[A]	1150
[B]	540
[C]	548
[D]	543
[E]	533
[F]	521
[G]	530
[H]	531
[I]	540
Status (= class label)	1151

Based on Table 3, one can immediately conclude that the data quality is poor due to a significant amount of lacking data points for some attributes. This conclusion was thoroughly analysed in further detail and the analysis is divided into separate analyses. This distinction is based on the earlier mentioned GC update resulting in two datasets (dataset *before GC update 2020* and dataset *after GC update 2020*) due to the major difference in data quality. This resulted in a dataset of 694 projects from the dataset *before GC update 2020* and 457 projects from the dataset *after GC update 2020*. Table 4 shows the populated records per attribute and the significant difference in data quality comparing the datasets *before GC update 2020* and *after GC update 2020*. The attributes within the dataset *after GC update 2020* have approximately 95-100% populated records, implying much better data quality.

Table 4. Overview of data quality per dataset

Attribute	**Populated records *before GC update 2020* (in number of rows)**	**Populated records *before GC update 2020* (in percentage)**	**Populated records *after GC update 2020* (in number of rows)**	**Populated records *after GC update 2020* (in percentage)**
[A]	694	100.00%	456	99.78%
[B]	97	13.98%	443	96.94%
[C]	101	14.55%	447	97.81%
[D]	101	14.55%	442	96.72%
[E]	92	13.26%	441	96.50%
[F]	87	12.54%	434	94.97%
[G]	91	13.11%	439	96.06%
[H]	91	13.11%	440	96.28%
[I]	89	12.82%	451	98.69%
Status (= class label)	694	100.00%	457	100.00%

Based on the information provided in Table 4, Figures 15 and 16 show the differences in data quality between both datasets. The figures show the average percentage of populated records and the average percentage of missing values. This is a major data challenge that is tackled along with the data description part, but the upcoming Data Preparation sections explain more data challenges.

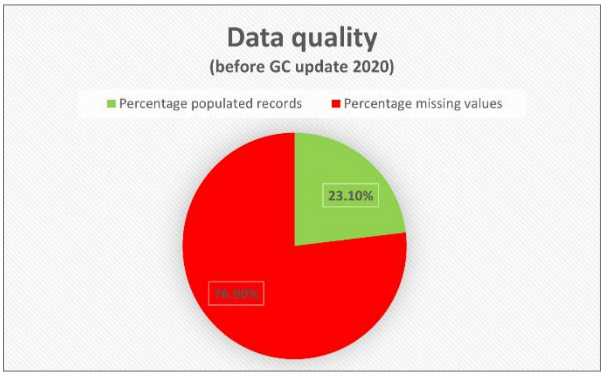


Figure 15. Data quality dataset *before GC update 2020*

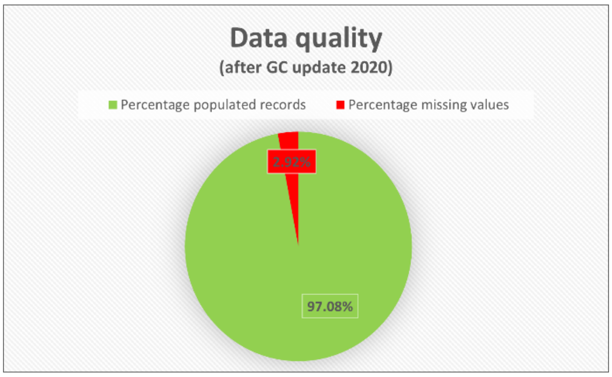


Figure 16. Data quality dataset *after GC update 2020*

4.1.3.2 Data Preparation

The second and last CRISP-DM data analysis phase involved in this study was the Data Preparation phase. This phase was about producing a dataset that is ready to use for Van Aken's Solution Design phase based on raw data. The raw data was converted into a consistent format to ensure the easiest possible way to get interpreted by a model. The output of the Data Preparation phase was a final created dataset and its description, which is selected, cleaned, constructed, integrated, and formatted (Czvetkó et al., 2022). This dataset was prepared in MS Excel and formatted to CSV (comma delimited) as this can be read by Python. Next, Python used this dataset in the code to establish a DT. The Data Preparation phase is divided into multiple sub-sections to provide a complete understanding of the data. The fundamental aspect of the sub-sections is data cleaning and comprises tokenization, filtering, and outliers, which are separately discussed below.

4.1.3.2.1 Tokenization

Tokenization is substituting a sensitive data element with a non-sensitive replacement (Singh & Manure, 2020). For example, the binary attributes (which means two possible outcomes e.g., yes and no) have multiple alternatives for the same outcome. Table 5 shows the tokenized elements, where the left column presents the used terminology for all the same outcomes: yes. Therefore, the tokenized description results in the right column of the table: again all yes, but yet better interpretable by a model. Additionally, all outcomes were translated to a numeric value so the yes' are 1 and the no's are 0. Besides the binary attributes, the other attributes have also a numeric replacement that is as easy as possible to interpret by the model. Due to confidentiality, it is not possible to explicitly state the replacements but the substitutes are three times formatted to a numerical scale of 1-9, so textual parts are removed. According to this substitution, duplicates in the dataset (e.g., same meaning but different definitions) were combined into one substitute.

Table 5. Example of tokenization binary attributes

Description in dataset	**Tokenized description in dataset**
'Yes'	'Yes'
'Yes, please specify'	'Yes'
'Yes, please specify which one(s)'	'Yes'

4.1.3.2.2 Filtering

Another data challenge is filtering out missing values. As discussed in section 4.1.3.1.2, the dataset *before GC update 2020* has many missing values resulting in poor data quality. However, the dataset *after GC update 2020* has only a few missing values and these missing values are negligible. All occurring missing values were deleted in the dataset as the open innovation projects are unique and substituting is impossible. The six models, three models per sub-dataset, resulted in a different number of missing values and the number of rows left in the dataset. Table 6 shows these results per model.

Table 6. Overview of missing values per dataset per model

Dataset	Model	Attributes	Number of missing values	Number of rows left
Before GC update 2020	1	[A]+[B]+[F]+[G]+[H]+[I]	3105	77
Before GC update 2020	2	[B]+[F]+[G]+[H]+[I]	3015	87
Before GC update 2020	3	[F]+[G]+[H]+[I]	2418	87
After GC update 2020	1	[A]+[B]+[C]+[D]+[E]+[F]+[G]+[H]+[I]	208	348
After GC update 2020	2	[B]+[C]+[D]+[E]+[F]+[G]+[H]+[I]	119	432
After GC update 2020	3	[C]+[D]+[E]+[F]+[G]+[H]+[I]	105	432

4.1.3.2.3 Outliers and inconsistencies

Finally, the last data challenge is capturing outliers and inconsistencies in the dataset. The outliers and inconsistencies were checked per attribute type (binary and not binary). First, the binary attributes were checked on outliers and inconsistencies. The two distributions of the binary attributes per dataset are presented in Figures 17 and 18. This analysis aimed to check whether the binary attributes are representative to include in the dataset because an attribute that has 99% yes' and 1% no's, the attribute is not representative. To investigate whether the attribute was included or not, a variance threshold of <25% was used. Hence, in all attributes where the yes category or the no category has less than 25% of the total samples in its category, the attribute was excluded from the dataset. Based on Figures 17 and 18, one can immediately conclude that the dataset *before GC update 2020* returns two attributes that cannot be selected as both attribute [D] and attribute [E] has less than 25% for the no category. On the contrary, all binary attributes remained in the dataset *after GC update 2020* because the threshold did not exclude an attribute.

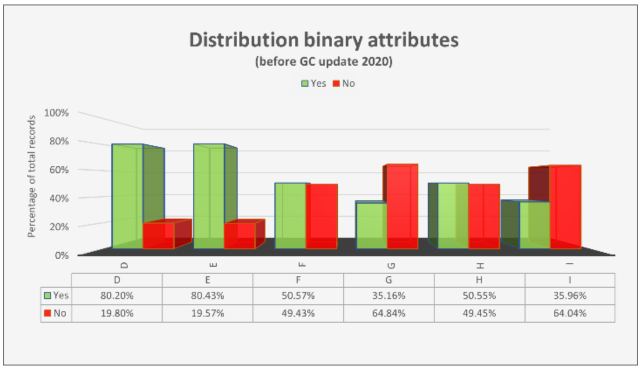


Figure 17. Distribution binary variables *before GC update 2020*

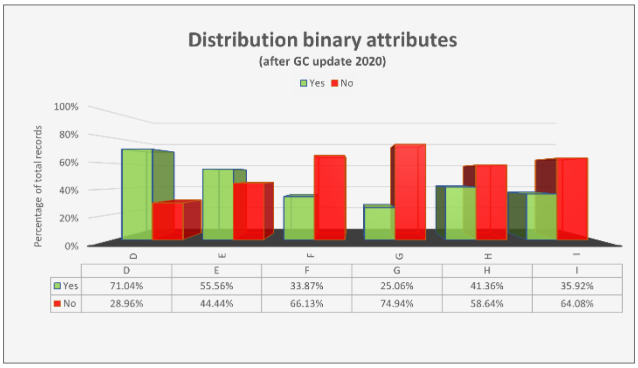


Figure 18. Distribution binary variables *after GC update 2020*

Moreover, the other attributes were checked to investigate whether this attribute is relevant and representative to include in the dataset. Figures 19, 20, and 21 show the distribution of attributes [A], [B], and [C]. First, although a major difference in data quality between the dataset *before GC update 2020* and *after GC update 2020*, attribute [B] has reasonable similarities. Therefore, this attribute is reliable and was taken into the dataset to create a final model. Next, contrary to the previously described attribute, attribute [C] is not representative and thus unreliable for the dataset *before GC update 2020*. The reason for this is the deviating number of records, which is inexplicable but the argumentation must remain anonymous due to confidentiality reasons. Finally, attribute [A] has many alternatives possible so threshold outliers are eliminated to get the correct understanding of the attribute. An attribute [A] is an outlier if attribute [A] contributes less than 1% to the dataset *before GC update 2020* or the dataset *after GC update 2020*. So, in the end, only attribute [A] was included if it contributes to both datasets equally or more than 1%. As can be seen in Figure 19, attribute [A] has also reasonable similarities between both datasets and, therefore, this attribute is representative and thus reliable for both datasets to include to create a possible final model.

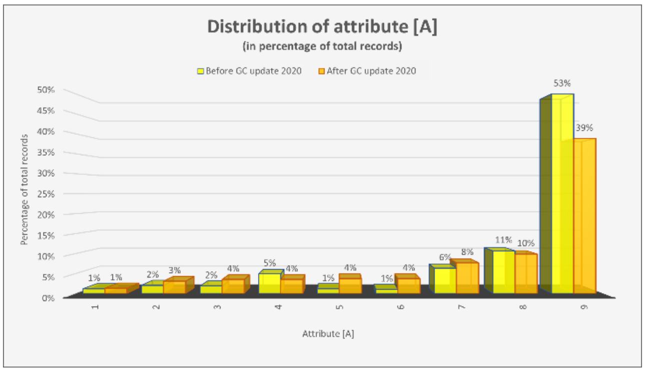


Figure 19. Distribution attribute [A]

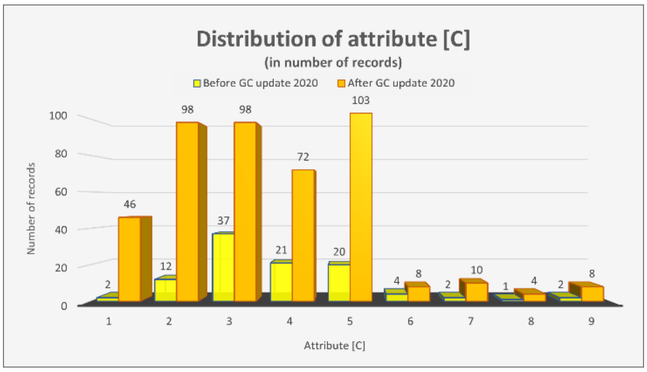


Figure 20. Distribution attribute [C]

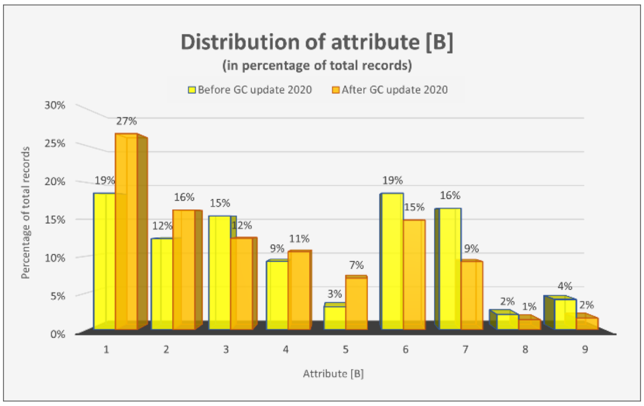


Figure 21. Distribution attribute [B]

4.1.4 Opportunities and limitations

This section of the Analysis & Diagnosis phase discusses the opportunities and limitations in the current process. Whereas opportunities imply potential areas for improvement that affects positively, limitations are so-called challenges that are most likely to disrupt the process and affect it negatively. First, the opportunities within the process converge to one major opportunity which arises to be the possibility to improve decision-making capability by deriving decision logic at decision points based on historical data. Currently, it is unknown what the decision-making capability is and historical data is not utilized while this valuable historical data is available for thorough analysis by ML algorithms. On the contrary, the current process explores limitations that will be either supported by the proposed solution design or proposed as a recommendation for further research due to limited resources. This section shortly explores the challenges that have the potential to get supported by the proposed solution design and section 5.3 describes the challenges proposed for further research as recommendations. The limitations are formulated based on the evidence collected, which is described in the corresponding methodology sections 3.2 and 3.3.

First, the process deals with a prioritizing problem as it is difficult to prioritize which projects to select for further analysis. This is especially a problem because of scarce resources implying limited time, people, and budget. Next, the terminology used in the process flow deviates due to lacking guidance in-process models and is therefore hard to understand. There is no clear visualization and explanation of the process flow with returning terminology in other documents. As a result of this challenge, it was uncovered that there is a biased way of approaching along with the process flow. There are multiple experts and it deviates whether an expert uses terminology and explains the flow. This indicates biased ways of executing the process flow not only in terminology but also in way of approaching it. Finally, data-related challenges are uncovered but are already discussed in section 4.1.3 about data-related analyses. In summary, inconsistent data registration was uncovered relating to incomplete information referring to the dataset *before GC update 2020*. It turned out that this is partly captured as since 2020 most fields in the submission form are made mandatory so the dataset *after GC update 2020* is reliable and the challenge does not hold for this dataset.

4.1.5 Conclusion Analysis & Diagnosis

In summary, this section discussed the current design of the open innovation process at Shell and the opportunities and limitations within this process. Previous to these elements, thorough data analysis has been carried out using the Data Understanding and Data Preparation phases of the CRISP-DM framework. Moreover, an extended analysis concerning data quality is done to get a detailed understanding of the data to prevent biased or useless data. All these analyses of the Analysis & Diagnosis phase are the foundation for the upcoming section about the proposed Solution Design where the combined ML and DMN approach for knowledge-intensive decision-making within open innovation is the core element.

4.2 Solution Design

After gaining knowledge about the current situation of the process and the opportunities and limitations in this process, this phase proposes a redesign as a possible solution to the problem statement. This section refers to step 6 of Figure 9 and comprises the Solution Design phase of the Van Aken framework. This phase is related to sub-question three: '*How can decision-making capability in Shell's open innovation process be improved?*'. This section uses multiple sections to give a complete overview of the Solution Design and consists of an introduction, the design itself, and a conclusion.

4.2.1 Introduction

The previous section, section 4.1, formed a starting point for this Solution Design section. The current design presented in section 4.1.2 could be interpreted as the design where the problem still holds, however, the redesign offered a solution approach for the problem stated. This section aims to create a redesign that improves decision-making capability, which refers to the ability to support decisions based on historical insights (Ghattas et al., 2014). Therefore, a combination of Machine Learning (ML) and Decision Model and Notation (DMN) was used because ML derives decision logic at decision points via decision-mining and DMN represents mined decision logic for visualization purposes. The ML part is the first model that served as a decision model and the DMN part is the second model that served as the expert knowledge model. Within this combination of models, the decision model is a filter model for the expert knowledge model. Using the decision model (based on an ML's Decision Tree), projects could be rejected in an earlier stage before going into the expert knowledge model (based on DMN modelling). The decision model is based on historical data and the expert knowledge model requires knowledge from Shell's experts that are required for the process because of its knowledge intensity.

4.2.2 Design

This section elaborates on the proposed redesign where the ML model serves as a filter model for the expert knowledge model. The first sub-section discusses the ML model and the second sub-section represents the mined decision logic by the ML model using DMN modelling.

4.2.2.1 Model

After understanding and preparing the dataset in section 4.1.3, six models were analysed to investigate what the best model is. However, it was already found that only three models have the potential to become a model for the design. These three models were based on three datasets extracted from the dataset *after GC update 2020*. It was found that the other three models based on the dataset *before GC update 2020* are not reliable due to poor data quality so conclusions can be hardly drawn. Table 7 presents the six models consisting of the dataset from which the model is extracted, the number of the model for the particular dataset, and the composition of attributes. Additionally, the accuracies of the DTs based on the three models extracted from the dataset *after GC update 2020* are shown because these models had the potential to integrate as earlier discussed.

Table 7. Overview of possible models

Dataset	Model	Attributes	Accuracy DT
After GC update 2020	1	[A]+[B]+[C]+[D]+[E]+[F]+[G]+[H]+[I]	68%
After GC update 2020	2	[B]+[C]+[D]+[E]+[F]+[G]+[H]+[I]	80%
After GC update 2020	3	[C]+[D]+[E]+[F]+[G]+[H]+[I]	74%
Before GC update 2020	1	[A]+[B]+[F]+[G]+[H]+[I]	-
Before GC update 2020	2	[B]+[F]+[G]+[H]+[I]	-
Before GC update 2020	3	[F]+[G]+[H]+[I]	-

As can be seen in Table 7, the best model is the second model extracted from the dataset *after GC update 2020* because this model returns the highest accuracy and, therefore, most powerful insights. This section elaborates on this best model found as this model was chosen to integrate as the proposed solution design for improving decision-making capability within the open innovation process at Shell. First, the input parameters of the ML model will be discussed, which are followed by the ML outcomes.

4.2.2.1.1 Input parameters

The dataset has two different sets, namely a training set to train the DT and a test set to test the DT's performance. These two sets were explicitly separated and the test set contains data that is never seen before by the model to prevent biased performances, which is set at 0.25 (25%). The performance of a DT could be improved or disrupted by tuning parameters, which was done by input parameters and this section discusses these parameters. Measuring the performance of the DT is central to the next section to test whether the DT gives meaningful insights. The best model consists of eight out of nine relevant attributes, which are attribute [B], attribute [C], attribute [D], attribute [E], attribute [F], attribute [G], attribute [H], and attribute [I]. It was found that attribute [A] disrupts the accuracy and therefore it was excluded. The attributes are so-called variables that affect a given outcome, or so-called label, that the model wants to predict. The eight attributes are the input, which are characteristics of the submissions via the GC portal, and the output or so-called label is whether the project is converted or rejected. A converted project is successful as the project moved to the [N] phase, and a rejected project is unsuccessful as the project left the system. Therefore, the label of the DT consists of two classes, namely class zero (0) representing rejected projects and class one (1) representing converted projects.

Many parameters were set and tuned to successfully create the final DT. The core parameter is the criterium which was set as entropy. Entropy is about measuring disorder, which can have multiple values but the extremes are zero (lowest) and one (highest). Whereas an entropy of zero means a pure split resulting in no disorder, an entropy of one means an even split (so-called fifty-fifty) resulting in an extreme disorder and no majority within the class (Bramer, 2007). The final features were chosen to set to the eight beforementioned attributes. Moreover, the DT has many parameters to tune such as the minimum samples leaf that defines the minimum number of samples required to be at a DT's leaf node. This input parameter was set at 22 because this represents around 5% of the total samples. Hence, the DT only returned leaves that had at least 22 samples in them. Another input parameter is maximum depth, which refers to the length of the DT's path from top to bottom. This was not necessary to use because the previously described minimum samples leaf parameter already shortened the DT's path.

4.2.2.1.2 Outcomes

The goal was to identify insights into which characteristics belong to successful projects and which characteristics belong to unsuccessful projects by deriving decision logic at decision points using ML. Table 8 shows the outcomes of the DT as an ML model, which shows only the valuable insights of the ML model. These insights are greatly supported by a significant number of projects. The complete dataset consists of 457 projects, the dataset after data cleaning (filtering on missing values) consists of 432 projects, and the test set consists of 108 (25% of 432) projects. The characteristics belonging to unsuccessful projects are given, however, the characteristics belonging to successful projects are not given. This is explained in the next section which takes a closer look at the evaluation of the solution design in terms of performance and validation. The implications of the outcomes of the ML model are further explained as part of the embedding in the DMN modelling approach in the next section.

Table 8. Outcomes of ML model

attribute [B]	**attribute [C]**	**attribute [D]**	**attribute [E]**	**attribute [F]**	**attribute [G]**	**attribute [H]**	**attribute [I]**	**Output**	**Support (# projects)**
-	1	YES	-	-	NO	-	-	Always reject	21
-	2	YES	-	-	NO	-	YES	Always reject	18
-	≥7	-	-	-	-	-	-	Always reject	10+4+8
-	≤2	NO	-	-	-	NO	-	Most likely reject	44 (42 vs 2)
≥5	5	-	YES	NO	-	-	-	Most likely reject	16 (15 vs 1)
1	≤3	-	NO	-	-	-	NO	Most likely reject	38 (37 vs 1)
1	≥4	-	-	-	-	-	-	Most likely reject	34 (33 vs 1)
1	-	NO	NO	-	-	-	-	Most likely reject	32 (30 vs 2)
1	-	YES	NO	-	-	-	-	Most likely reject	47 (46 vs 1)

4.2.2.2 Process

The ML outcomes represent the value-added paths of the DT in the form of a row in the table presenting decision logic. These outcomes indicate the characteristics belonging to unsuccessful projects based on historical data. As a result, the combination of particular characteristics was combined with an automatically generated output by the ML model to either always reject or most likely reject a project. Nine different combinations of characteristics, which were greatly supported by a significant number of projects, were derived from the event log consisting of historical data and can be seen in Table 8. The first three insights imply always rejected projects, so the output is to reject projects with these characteristics: *attribute [C] (1) + attribute [D] (YES) + attribute [G] (NO)*, *attribute [C] (2) + attribute [D] (YES) + attribute [G] (NO) + attribute [I] (YES)*, and *attribute [C] (≥7)*. The remaining six insights imply most likely rejected projects, where the output is to most certainly reject projects with these combinations of characteristics: *attribute [C] (≤2) + attribute [D] (NO) + attribute [H] (NO), attribute [B] (≥5) + attribute [C] (5) + attribute [E] (YES) + attribute [F] (NO)*, *attribute [B] (1) + attribute [C] (≤3) attribute [E] (NO) + attribute [I] (NO)*, *attribute [B] (1) + attribute [C] (≥4)*, *attribute [B] (1) + attribute [D] (NO) + attribute [E] (NO)*, and *attribute [B] (1) + attribute [D] (YES) + attribute [E] (NO)*.

The output of most likely reject projects rather than always reject projects is based on the fact that the ML model did not return an entropy of 0. This indicates that there is no pureness, i.e. the majority of rejected projects are disturbed by some converted projects. For example, the combination of *attribute [B] (1) + attribute [D] (YES) + attribute [E] (NO)* returns 46 rejected projects out of a total of 47 projects but there is still 1 converted project based on the same combination of characteristics. Therefore, the combination of these characteristics is supported by the output to most likely reject rather than always reject. These automatically derived outputs by the ML model were embedded within the BPMN modelling approach, which can be seen in Figure 22. Figure 22 shows the adjusted BPMN model, which represents the redesigned process flow. The explanations of the objects used in the BPMN model can be seen in Appendix B. The yellow box is the added activity that is about making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. This refers to the ML outcomes in Table 8 as underlying decision specifications as a base for experts' decisions.

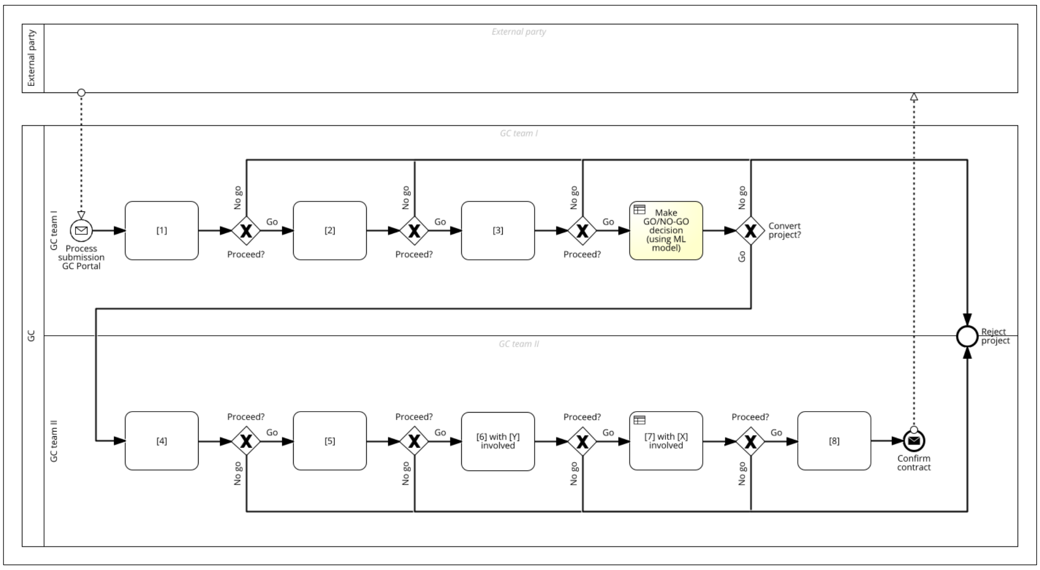


Figure 22. Redesign via BPMN

The added activity is an additional supportive stage for Shell's experts to base their future decisions on supportive and historical insights based on historical data. This new approach improves decision-making capability as historical data is used to derive decision logic at decision points that ensure the ability to support decisions based on historical insights. However, the BPMN modelling approach is just the visualization of the updated process design. Figure 23 shows the underlying decision logic visualized by decision specification in a DRD as the top level of the DMN modelling approach. The added activity is considered semi-automatic as the automatically derived outputs from the ML model resulting in the DMN decision table serves as advice to Shell's experts. The part of deriving decision logic at decision points is fully automated and refers to the approval rules using the ML model. However, the part of using these supportive insights to base decisions is manually done by Shell's experts and refers to the determination of approval implying to make a *GO* or *NO-GO* decision. The reason is that the new approach has a supportive nature by simplifying the work of Shell's experts rather than automatically replacing their work as the knowledge of Shell's experts remains the crucial and leading factor. The input data elements of the decision are the eight relevant attributes of the DT. The explanations of the objects used in the DMN model can be seen in Appendix B.

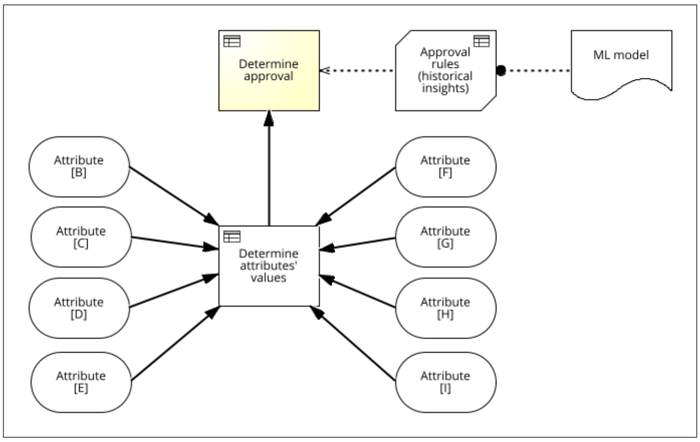


Figure 23. Decision model visualized via DMN DRD

As explained in Chapter 2, the theoretical background, the second level of the DMN modelling approach is the decision logic level that is commonly presented using a decision table representing decision rules. This embedded decision logic level in Signavio (DMN builder) can be seen in Figure 24 and is equal to the ML outcomes presented in Table 8. The ML model discussed in section 4.2.2.1 is the decision model that serves as a filter model for the expert knowledge model. This expert knowledge model is the most crucial expert moment in the process as explained in section 4.1.2 and visualized in Figure 12.

	Inputs																Outputs	Annotations
U	[redacted]																	
1	=	true	O	is not defined	=	1	=	false	O	is not defined	O	is not defined	O	is not defined	O	is not defined	"Always reject"	21
2	=	true	=	true	=	2	=	false	O	is not defined	O	is not defined	O	is not defined	O	is not defined	"Always reject"	18
3	O	is not defined	O	is not defined	≥	7	O	is not defined	O	is not defined	O	is not defined	O	is not defined	O	is not defined	"Always reject"	22
4	=	false	O	is not defined	≤	2	O	is not defined	O	is not defined	=	false	O	is not defined	O	is not defined	"Most likely reject"	44 (42 vs 2)
5	O	is not defined	O	is not defined	=	5	O	is not defined	=	true	O	is not defined	=	false	≥	5	"Most likely reject"	16 (15 vs 1)
6	O	is not defined	=	false	≤	3	O	is not defined	=	false	O	is not defined	O	is not defined	=	1	"Most likely reject"	38 (37 vs 1)
7	O	is not defined	O	is not defined	≥	4	O	is not defined	O	is not defined	O	is not defined	O	is not defined	=	1	"Most likely reject"	34 (33 vs 1)
8	=	false	O	is not defined	O	is not defined	O	is not defined	=	false	O	is not defined	O	is not defined	=	1	"Most likely reject"	32 (30 vs 2)
9	=	true	O	is not defined	O	is not defined	O	is not defined	=	false	O	is not defined	O	is not defined	=	1	"Most likely reject"	47 (46 vs 1)

Figure 24. Underlying decision rules via DMN decision logic level

4.2.3 Conclusion Solution Design

In summary, this section showed the combination of ML DT and BPMN + DMN as a proposed solution design for improving decision-making capability within Shell's open innovation process. Ideally, both insights into characteristics belonging to successful projects and unsuccessful projects are derived from the ML model. However, the DT as ML model only gives insights into derived decision logic at decision points for the rejected class representing unsuccessful projects. The performance analysis in the next section gives additional insights and explanations about this first impression. However, the ML outcomes for the rejected class do give powerful insights and are supported by a significant number of projects. The biggest change in the redesigned process flow is the additional activity of making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. This activity is supported by the DMN possibilities of presenting underlying decision logic clearly and understandably, which can be easily visualized. These supportive insights make the process more transparent and improve decision-making capability as historical data is used and future decisions have better pillars. The remaining part is the evaluation phase, which focuses on the value of the models for meeting the business goals defined at the preliminary stage of the case study research methodology. This evaluation is central to the upcoming section.

4.3 Evaluation

This section refers to step 7 of Figure 9 and comprises the Evaluation phase of the Van Aken framework, where the goal is to evaluate the model's results compared to the business value (Van Aken et al., 2012). This part focuses on the value of the models for meeting the business goals defined at the preliminary stage of the case study research methodology and is related to sub-question four: '*How can the proposed solution design be evaluated?*'. This section uses multiple sections to give a complete overview of the Evaluation phase and consists of an introduction, performance, validation, and a conclusion.

4.3.1 Introduction

It is crucial to investigate whether the model's results add value to the knowledge-intensive decision-making process within open innovation. This is a difficult part as insights into performances of decision-making capability are difficult to gather. This section aims to combine both a quantitative and qualitative evaluation method of the new model's performance. The quantitative evaluation method is performance measurement using ML model evaluation methods. Typically, accuracy is a well-known term that is used for performance measurement. This was used to investigate whether the outcomes are reliable and the predictions can be fulfilled. Next, the qualitative evaluation method equals the only possible validation method. Expert knowledge is the only option to use where Shell's experts are the core value. This expert knowledge is in addition to the model's results to validate that the model is an improvement. The study's results should improve decision-making capability. The evaluation aims to investigate to which extent the study's results are meaningful and to be able to conclude to which extent the combined approach of ML and DMN modelling is satisfactory within open innovation.

4.3.2 Performance

This section discusses the performance of the proposed solution design presented in the previous section. The performance of a DT is improved or disrupted by tuning parameters, which is done by input parameters as described in the previous section. The upcoming first sub-section discusses the performance of the ML model and the second sub-section elaborates on the process performance as a result of the new combined approach of ML and DMN for knowledge-intensive decision-making.

4.3.2.1 Model

As discussed in Chapter 2, theoretical background, one of the benefits of ML algorithms is that they are relatively easy to evaluate. A classification report is a well-known ML evaluation method, which is highly readable and understandable and consists of four commonly used metrics: accuracy, precision, recall, and F1-score. The equations of these four commonly used metrics for evaluating ML algorithms are shown below in Equations 1-4 (Zineb et al., 2021).

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad \text{(Equation 1)}$$

$$Precision = \frac{TP}{TP + FP} \quad \text{(Equation 2)}$$

$$Recall = \frac{TP}{TP + FN} \quad \text{(Equation 3)}$$

$$F1\ score = \frac{2 * Precision * Recall}{Precision + Recall} \quad \text{(Equation 4)}$$

These four equations rely on a matrix resulting in four types of values: True Negatives (TN), False Negatives (FN), False Positives (FP), and True Positives (TP). The distinction is made based on two segments: predicted values, which are predicted by the model, and real values, which are the actual values. Both predicted values and real values correspond to two options, namely an option for class 0 (no – rejected projects) and class 1 (yes – converted projects). Based on the explanation of Verma et al. (2022), all types will be briefly explained. TN represents a value that is both negative for the predicted and real value. Next, a value is FN when the predicted value is negative but the real value is positive. If reversed, i.e. the predicted value is positive and the real value is negative, the value is FP. Finally, when both the predicted and real value are positive, the value is TP. Table 9 shows this matrix, based on Verma et al. (2022), resulting in the four types of values used for Equations 1-4.

Table 9. Matrix of types of values

		Predicted values	
		Class 0	**Class 1**
Real values	**Class 0**	True Negatives (TN)	False Positives (FP)
	Class 1	False Negatives (FN)	True Positives (TP)

Based on the intuition behind Table 9, the four metrics of the classification report, as shown in Equations 1-4, are derived. The accuracy returns the proportion of the correct predictions, and the F1-score is the harmonic mean of the precision and recall scores. Whereas the precision returns the proportion of the TP among the total positive predicted values, the recall returns the proportion of the total correctly positive values predicted. Besides the four commonly used metrics, the classification report consists of three more elements, which are support, macro average, and weighted average. First, support is the number of observations in the particular class. Next, the macro average is the average between the two possible classes (class 0 or class 1 in this study), and the weighted average is a weighted average where the weights are the number of observations for the particular classes.

Table 10 presents the classification report, and thus the performances, of the best model that is integrated into the proposed solution design. This classification report is created using Python's module Scikit-learn. Based on this dataset and model, the combined approach of ML and DMN is created in the open innovation context. Table 10 shows that the accuracy is 80%, which is good accuracy. Particularly, this is not even good accuracy but also ideal and realistic accuracy. This is realistic because 100% accuracy is not good in ML theories. The reason is that this probably means that the algorithm is tested on the same data as the algorithm is trained.

Table 10. Classification report of best model (model 2)

	Precision	Recall	F1-score	Support
Class 0	0.82	0.95	0.88	85
Class 1	0.56	0.22	0.31	23
Accuracy			0.80	108
Macro average	0.69	0.59	0.60	108
Weighted average	0.76	0.80	0.76	108

Moreover, the classification report returned that there is a significant distinction in meaningful insights in rejected projects and converted projects. Deriving characteristics that belong to successful projects turned out difficult and the DT did not result in this part of the desired decision logic at decision points. The ML model only gives insights into derived decision logic at decision points for class 0 representing rejected (unsuccessful) projects. Again, precision means how good the model is at predicting a class, which is 82% for rejected projects and 56% for converted projects. Moreover, the recall metric returns a significant difference in explaining how many times the model was able to detect a class, which is 95% for rejected projects and only 22% for converted projects. This explains why the ML model only gives valuable insights into derived decision logic at decision points for the rejected class representing unsuccessful projects. The same holds for the other possible models (model 1 and model 3). However, these models have even worse precision and recall scores and likewise a worse accuracy. Therefore, these models are not relevant to select, which supports the decision to choose and integrate model 2. The classification reports of these unselected models can be seen in Tables 11 and 12.

Table 11. Classification report of other possible model (model 1)

	Precision	Recall	F1-score	Support
Class 0	0.80	0.78	0.79	68
Class 1	0.29	0.32	0.30	19
Accuracy			0.68	87
Macro average	0.54	0.55	0.55	87
Weighted average	0.69	0.68	0.68	87

Table 12. Classification report of other possible model (model 3)

	Precision	Recall	F1-score	Support
Class 0	0.80	0.88	0.84	84
Class 1	0.38	0.25	0.30	24
Accuracy			0.74	108
Macro average	0.59	0.57	0.57	108
Weighted average	0.71	0.74	0.72	108

As opposed to the good performance of the best model chosen to integrate as redesign, Table 13 shows the results of the classification report of one of the three invalid models due to poor data quality extracted from the dataset *before GC update 2020*. This additional part aims to show that the data quality in the final chosen model is particularly good compared to the other models, especially the models extracted from the dataset *before GC update 2020*. Although this model is the best model of these three invalid models based on datasets extracted from the dataset *before GC update 2020*, the accuracy is poor as can be seen in Table 13. This is due to the poorly populated dataset based on only a few data points (number of rows). Table 13 confirms the unreliable and poor models from the dataset *before GC update 2020*. Therefore, the dataset *after GC update 2020*, and especially model 2 as the best model chosen, is reliable and the DT has good accuracy.

Table 13. Classification report of worst model

	Precision	**Recall**	**F1-score**	**Support**
Class 0	0.11	0.17	0.13	6
Class 1	0.67	0.56	0.61	18
Accuracy			0.46	24
Macro average	0.39	0.36	0.37	24
Weighted average	0.53	0.46	0.49	24

4.3.2.2 Process

Additional to the previous section about measuring the performance of the ML model, this section elaborates on the performance of the process. The biggest change in the redesigned process flow is the additional activity of making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. This adjusted process flow, called the redesign, is supported by an activity that relies on both the ML model and DMN's decision-mining possibility to visualize underlying decision logic clearly and understandably. The core of the added activity provides a base for making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. This activity has a supportive nature by supporting and simplifying the work of Shell's experts required rather than a replacement nature as input of Shell's experts remains the crucial and leading factor. These insights make the process smarter, which ensures more transparency in both the process and decision-making. Therefore decision-making capability is improved and historical data is used to derive these insights to provide better pillars for future decisions.

Moreover, the ML model, based on data extracted in the [M] phase, serves as a filter model for the expert knowledge model in the [N] phase. As open innovation projects are of high strategic importance where many resources are involved, resources should be allocated to the most valuable projects. With the new approach, unsuccessful projects are less likely to be successful so the resources will be more allocated to the valuable or so-called successful projects. Therefore, Shell's experts have prior knowledge about (un)successful projects so experts' time can be used more effectively. Currently, the existing process and data available have no parameters to measure performance. Therefore, it is difficult to express process performances in terms of resource savings in numbers. It is unknown how much time Shell's experts spend identifying and evaluating projects accordingly. These crucial data points to calculate concrete savings are missing and are a practical limitation of the study. Due to this study's timeframe, it is impossible to discover these savings, which is a recommendation for further research. However, it is possible to compare the decision outcomes of the existing approach to the decision outcomes of the new approach. The predicted decision outcomes of the new approach are measured by the confusion matrix evaluation method. A confusion matrix is a matrix where the columns and rows equal the numbers of classes, and its intuition exactly relates to the intuition behind Table 9 (Sachan et al., 2020). The confusion matrix was created using Python's module Scikit-learn, and the confusion matrix of the integrated model can be seen in Table 14. The FP instance and FN instance are the instances that are incorrectly predicted, but these instances did not have major influences as the total number of projects is only 20 out of 108, which relates to the accuracy of approximately 80%.

Table 14. Confusion matrix of best model (model 2)

		Predicted values	
		Class 0	**Class 1**
Real values	**Class 0**	TN = 86	FP = 3
	Class 1	FN = 17	TP = 2

Due to 80% accuracy, the predicted values are not the same as the real values. The existing decision outcomes were presented in section 4.1.2 and Figure 25 presents this again for easy comparison purposes. Figure 26 presents the decision outcomes based on the new approach, which are based on the measurement by the confusion matrix. As can be seen, there is a major difference in the percentage of rejected projects compared to converted projects as the new approach rejects almost all projects. The reason is that hardly any pattern could be found in the converted class due to its poor precision and recall scores and this class has a significantly lower number of projects available.

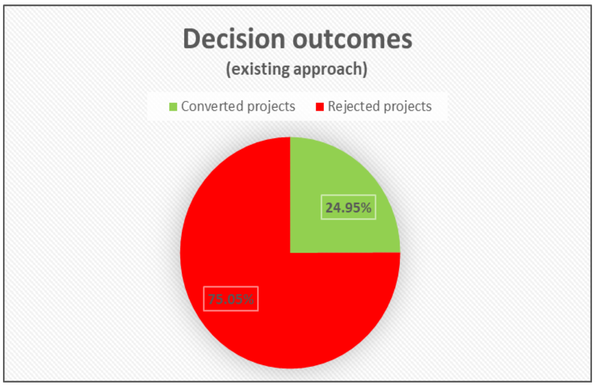


Figure 25. Decision outcomes existing approach

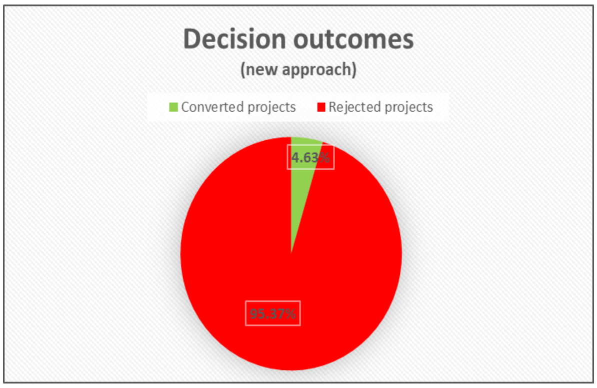


Figure 26. Decision outcomes new approach

4.3.3 Validation

The embedded ML model in the DMN modelling approach as a supportive process model to the current situation is a theoretical model, which is not currently testable in practice. However, it should be validated to compare the model's results to the business value. Therefore, expert knowledge is the proposed measurement to validate the redesign for initial steps toward future implementation. Expert knowledge is about discussions with stakeholders to gather the experts' perspectives. The goal of expert knowledge validation is to validate the correctness and effectiveness of the model. Moreover, it validates the supportive nature and the value of insights. The upcoming first sub-section discusses the ways of validation and the second sub-section shows the results of the validation method chosen.

4.3.3.1 Ways of validation

Expert knowledge can be used in different ways and at different stages of the research. The first stage is validating the information gathered in sub-question one and two to define the correct starting point. This is done in two different ways along with the study's timeframe. First, it is a self-directed validation by analysing documents and databases independently from Shell's experts to discover flows and challenges. In this way, validation is done to check the content gained from interviews with Shell's experts against the real situation. Next, feedback or information from interviews is validated by the experts themselves. Once input from experts was processed, the information gained is discussed with the particular expert and other experts as well to review the correctness of the information.

The second stage of validation is the validation of the final model created in sub-question three and can be done using expert knowledge only. As discussed in the introduction of this section, the content of the redesign should be checked by experts as this increases the reliability of the research even more. Many validation methods are possible, however, the methods differ in thoroughness and depth. The differences in thoroughness and depth depend on multiple factors such as sample size and type of format used, which results in a wide range of validation methods. This study uses a validation method where a project evaluation form is a basis. All experts have to fill in this evaluation form individually to prevent bias or influences among the experts. This results in more powerful feedback that is of complementing nature rather than overlapping nature. In addition, experts can take their time appropriately instead of quickly answering the question because of a joint meeting at an unfavourable time. The chosen validation method uses a Likert scale (ranking method from poor to excellent) as a basis to indicate an expert's opinion with additional opportunities to explain the scores or to raise awareness about other aspects that are not mentioned in the questions. Attached to the project evaluation form is a summary of the study and additional content for understandability. The next section explains the experts involved and the results of the beforementioned validation method.

4.3.3.2 Results validation

This section describes the results of the chosen validation method mentioned in the previous section. Four experts were involved in this research validation method with four different functions representing four different disciplines. All experts involved are stakeholders of Shell's open innovation process but with different backgrounds to approach the helicopter view rather than the centralized perspective. Due to confidentiality, the exact roles of the experts involved cannot be explicitly stated, but the four experts involved in this research represent four different high-level functions. Almost all experts got their PhD in their expertise field (mostly electrical/chemical engineering) with significant professional experiences. The format for the project evaluation form used as guidance for this validation method can be seen in Appendix C. The evaluation form consists of five questions that allow the experts involved to thoroughly understand the project. The five evaluation questions consist of the following elements: well-organized models, correspondence to practice, hands-on approach, old versus new model, and project's results versus business value. First, the results are given in Figure 27 and a more detailed discussion of these results follows per question in the remainder of this section.

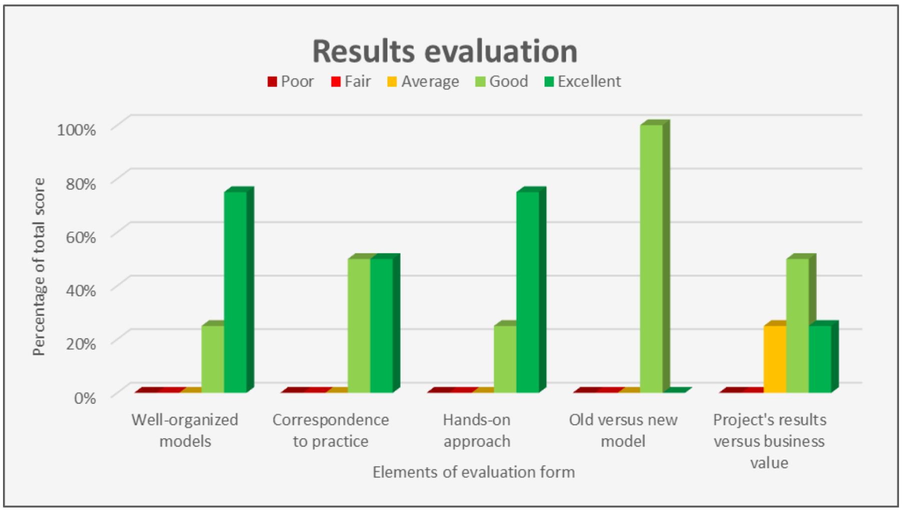


Figure 27. Results evaluation

4.3.3.2.1 Well-organized models

The first element that allowed the experts involved to thoroughly understand the project is the degree of well-organized models. In general, the experts indicated that the models are well-organized and even the majority indicated that the models are excellent in terms of well-organized. The experts specified that the models are very clear and stated that the contribution of the model to the decision process is illuminating. Next, the information flows, (decision logic at) decision points and explanations of the models are clear because of the practical usefulness of the models rather than using complicated mathematical models especially. To conclude, the complex and complicated set of processes is converted into highly readable models that experts particularly appoint useful in their organization.

4.3.3.2.2 Correspondence to practice

The second element that allowed the experts involved to thoroughly understand the project is the degree of correspondence to the practice. The degree of correspondence to the practice is equally good as well as excellent, which indicates that the study is close to the practice. The core element that ensures the high comparison to practice is the problem-worth-solving nature of the study. Although the models are not complicated mathematical models, the language and mathematics involved describe the situation correctly. To conclude, the methodologies are clear and well addressed and the application of the decision-making technologies is achieved in line with the practice.

4.3.3.2.3 Hands-on approach

After the elements of well-organized models and correspondence to the practice, the next element allowed the experts to have a thorough look at the project's degree of a hands-on approach. The majority indicated that the models are excellent in terms of a hands-on approach. Therefore, there is no lack of a hands-on approach and the experts summarized the work as practically relevant. One of the experts indicated that discussion and reflection points are missing, but this is captured by Chapter 5 of this report. The theoretical background and additional documents such as Appendix B attached to the project's results made the study easily understandable for a hands-on approach. As a result, there is no biased understanding due to deviating backgrounds of the experts.

4.3.3.2.4 Old versus new model

The fourth element that allowed the experts involved to thoroughly understand the project is the comparison of the new model to the old model. The new model is good compared to the old model, which is agreed upon by all experts. The new model is better than the existing or so-called old model and is even superior to what the experts are doing (or not) in the past. The experts indicated that the new model and its results are a huge learning experience that suggests ideas that have not been thought about before. The only thing that could be improved, which keeps from the excellent score, is the complexity of the problem in terms of opportunities to tackle it. The research is limited to a specific scope that could not deeply dive into all opportunities manifested in the old model. The experts did not indicate this as detracting from the current research but as an indictment of the organization's culture. Although the experts explicitly mentioned that the analysis, models, illustrations, and conclusions of the new model are valuable to use and build on, for the time being, the experts are curious to incorporate more data and approaches that go out of this research' scope. This is a desirable next step that could allow the organization to benefit even more from the new model's results in detail.

4.3.3.2.5 Project's results versus business value

Finally, after the previous elements, the concluding element that allowed the experts to have a thorough look at the project is the comparison of the project's results to the business value. The experts' scores on this element deviate most from each other, however, the majority indicated that the project's results versus the business value are good. Although the previous section indicated that the limited scope declined some opportunities manifested in the old model, the defined scope is very detailed and complete. The defined scope is very narrow and the experts summarized that the research deeply dives into symptoms, root causes, and potential solutions for the defined scope. The project's results are powerful and ensure gaps and possibilities for improving the decision process, and this reflection and growth in the specific domain is highly appreciated.

Although the project's results are highly rated and of value-adding nature, some aspects of the project's results could be better. First, the outstanding question is how more advanced models get valuable insights into the parameters influencing the decisions. Next, the timeframe of the research should have been longer to thoroughly make the adjustments and measure the positive impacts of changes made. As mentioned during this study, the implementation phase is lacking due to the timeframe and the dependency on larger-scale research. The opinion of the experts perfectly suits this missing phase and is one of the recommendations. To conclude, the research consists of a complex problem to solve, with poor data quality, but the project's results adapt and search for appropriately qualify for analysis. The project's results clarified the way of decision-making based on a structured overview of the process. Besides this, the experts stated that the research's data analysis parts revealed intriguing facts that would not have been aware of without the project.

4.3.4 Conclusion Evaluation

This evaluative section comprised both the proposed solution design's performances and results of the validation to completely evaluate the study and to discuss the practical relevance to the specific business case. The results of the proposed solution design lead to lots of thoughtful feedback points. Based on the performances, it can be concluded that the model's results are reliable and, therefore, the proposed solution design is realistic and feasible. This evaluative section focused on the value of the model for meeting the business goals defined at the preliminary stage of the case study research methodology, where the problem definition was one of the core elements. The problem statement of this study for the business case was: *the current open innovation process at Shell does not utilize historical data to gain insights into decision-making capability.*

Based on the model's results, the quantitative evaluation method, and both the qualitative and validation method, the lacking insights in the decision-making of the open innovation are supported by the solution design and are particularly reliable as both the performance and validation are good. The model's results are based on historical data and present decision logic at decision points indicating criteria. When using the combined ML and DMN approach, Shell's experts are relieved by a supportive decision model that filters, i.e. separate using the ML model, the (un)successful projects based on the project's characteristics. Hence, the process is smarter, which ensures more transparent decision-making and, therefore, an improved decision-making capability. Here, the study ends as there will be no implementation phase of the Van Aken framework due to the timeframe and the dependency on larger-scale research. However, the study has outstanding steps that could be further explored for both the business case and academic research. Chapter 5 discusses these further steps in the recommendations section.

4.4 Generalization

This finalizing section elaborates on the generalization part and refers to step 8 of Figure 9. As the results of this case study are context-specific to Shell, the generalization investigates to which extent the study's findings can be generalized to other contexts. Therefore, this section aims to draw conclusions from the particular study's context to a broader context and provide future research with key findings that might increase its effectiveness and efficiency. The reason to discuss this part separately from the study's conclusions is that this step is explicitly part of the case study design.

First, the business case focused on the GC process as one of the open innovation processes at Shell. Shell is a large company consisting of more open innovation processes, such as the Shell Technology Ventures (STV). The study's dataset only considered GC open innovation projects, which are projects related to the scope of GC implying technology developments that commonly comprise a Technology Readiness Level (TRL) of approximately 2, 3, 4 or 5. However, the problem of not utilizing historical data to gain insights into decision-making capability might also occur in other departments. The STV comprises another range of TRLs. Although the open innovation projects have other characteristics due to different ranges of TRLs, the approach can be applied in other departments, e.g. STV, that want to support the decision-making by analysing historical data to improve decision-making capability. Based on the methodology, it is expected that datasets can be easily created by extracting input attributes (characteristics) and assigning an output label as a base for the ML model. However, it is expected that the ML model only properly works once datasets are well prepared, where good data quality is an important dimension.

Next, the proposed combined ML and DMN approach can be applied to many other open innovation processes. It is foreseen that the methodology is suitable for other open innovation processes that desire to derive decision logic at decision points via decision-mining using ML and to represent the mined decision logic via DMN. For example, the input for the ML model is based on characteristics extracted from a dataset. These characteristics represent different types of variables, with different natures and variants. After scaling the data, the combination of binary attributes with other types of attributes (cannot be stated due to confidentiality) worked well in this study, and therefore, it is expected that a wide range of other characteristics can be selected as well in other studies. Because the characteristics (implying input characteristics for the ML model) are extracted from a dataset within this study's sample, it is expected that characteristics can also be extracted from other datasets outside this study's sample. Hence, the decision criteria can be identified to derive decision logic via decision mining. However, some assumptions should be made to properly use the approach in various contexts. First, the approach works well on well-prepared datasets consisting of good data quality, but these datasets cannot be too large. The efficiency of the approach works best on small data sets as it is required to avoid large amounts of data. Besides this assumption of the dataset, the open innovation process itself should have similar characteristics to the scoped innovation process where the approach is implemented using the frameworks of Van Aken and CRISP-DM. The nature of this assumption refers to the process flow where experts should make decisions on technology developments that are crucial for the organization. Therefore, it is expected that the open innovation process should need resources implying people (experts) whose time should be allocated best due to the high strategic importance. Based on this, the technology developments enter a knowledge-intensive programme where many stages should be completed to finally deliver a successful technology development. Moreover, the open innovation process should not utilize historical data in another way because this may cause complications in the process.

Today’s increasing amounts of data available in open innovation processes result in lots of potential other cases, and it is expected that this could be perfectly treated by this study’s approach. For example, the ML model resulted in 80% accuracy and the outcomes are supported by a significant number of projects. Therefore, it is expected that the ML model’s accuracy in other open innovation processes will be also good and realistic. Again, this expectation is under the condition of pre-screening the dataset and good data quality. However, the approach only successfully derived decision logic at decision points for the rejected class instead of both classes. Therefore, the performance of indicating success criteria rather than failure criteria is lacking. Because of this significant difference in successful projects compared to unsuccessful projects within this study’s sample, it is expected that this difference also occurs outside this study’s sample. Nevertheless, this expectation confirms a well-known conclusion of James Dyson that organizations should learn from these insights in failure as learning from success is rarely the case within organizations (Wulfen, 2016). In summary, under the conditions of pre-screening the dataset and good data quality as input, the combined ML and DMN approach can be generalized to various contexts with the assumptions and resources committed.

“Enjoy failure and learn from it. You never learn from success.”

- James Dyson -

5 Conclusion

This concluding chapter consists of multiple sections to discuss the overall conclusion and make recommendations based on the overall findings. This chapter discusses the study's conclusions, divided into the case study conclusions and research conclusions, and the limitations and recommendations for both the business case and future academic research.

5.1 Case study conclusions

This first concluding section describes the case study conclusions by answering the formulated sub-questions specific to the Shell case. These sub-questions were formulated specifically for the case at Shell to answer the main research question, which will be discussed in the next section accordingly. The case study was supported by four sub-questions about the current design, the opportunities and limitations within this current design, the redesign, and the evaluation of this redesign. First, this study started at the Analysis & Diagnosis phase of the Van Aken framework where the first two sub-questions were central, namely '*How is Shell's open innovation process currently designed?'* and '*What are the opportunities and limitations in Shell's open innovation process?'*. The current design is for the first time investigated and visualized as there are no powerful existing models for the open innovation process at Shell. The current design is presented via the combination of BPMN and DMN modelling approaches. To be able to create and visualize this current design, information was gathered via interviews with Shell's experts involved and validated through self-directed analyses using documents and databases independently from Shell's experts to discover flows and challenges. The most important distinction in the current design is the [M] phase versus the [N] phase, where the [M] phase considers the input for the decision model that serves as a filter for the expert knowledge model that is created for the [N] phase. The expert knowledge model comprises the determination of approval for an open innovation project by [X] as this is the most critical stage where expert knowledge is highly important. The quantitative part is fully gathered via GC's database. This data analysis part is thoroughly done and supported by two CRISP-DM phases to create a useful, unbiased, and interpretable dataset for a final model. The data analyses resulted in many data challenges such as a major difference in data quality that has led to a division of the dataset into two different datasets. Splitting the dataset into two different datasets was a well-fitting solution as the dataset *after GC update 2020* resulted in a well-prepared dataset consisting of good data quality. The combination of gathering information to create and visualize the current design and the data analyses were the fundamental base for creating a proposed solution.

The Analysis & Diagnosis phase formed a starting point for the Solution Design phase, where the redesign was central. The earlier mentioned design was the design where the problem still holds, however, the redesign offered a solution approach for the problem stated. The problem was that Shell's current open innovation process does not utilize historical data to gain insights into decision-making capability. Within the Solution Design phase the third sub-question was central, namely '*How can decision-making capability in Shell's open innovation process be improved?'*. The proposed solution design was the combination of ML DT and BPMN + DMN for improving decision-making capability within the open innovation process at Shell. The biggest change in the redesigned process flow is the additional activity of making a *GO* or *NO-GO* decision based on the project's characteristics using the ML model. These insights make the process more transparent and improve decision-making capability as historical data is used and future decisions have better pillars. Ideally, both insights into characteristics belonging to successful projects and unsuccessful projects are derived from the ML model. However, deriving characteristics that belong to successful projects turned out difficult and the

DT did not result in this part of the desired decision logic at decision points. The DT as ML model only gives insights into derived decision logic at decision points for the rejected class representing unsuccessful projects, which is supported by a significant number of projects so these are noticeable insights.

Additionally, the models are evaluated in the Evaluation phase that is based on the fourth and last sub-question, namely '*How can the proposed solution design be evaluated?'*. The models are evaluated using a quantitative and qualitative method comprising the proposed solution design's performances and the model's validation using expert knowledge. Based on the performances, it can be concluded that the model's results are reliable and, therefore, the proposed solution design is realistic and feasible. Particularly, the accuracy of the ML model was 80% but the model showed different performances for the unsuccessful projects compared to the successful projects as described earlier. The DT as ML model only gives insights into derived decision logic at decision points for class 0 representing rejected (unsuccessful) projects. It was found that it is difficult to find patterns in converted projects because all patterns showed up have approximately an entropy of 1, which means equally divided so 50% true for converted projects but also 50% true for rejected projects. Again, precision means how good the model is at predicting a class, which is 82% for rejected projects and 56% for converted projects. Moreover, the recall metric returns a significant difference in explaining how many times the model was able to detect a class, which is 95% for rejected projects and only 22% for converted projects. This explains why the ML model only gives valuable insights into derived decision logic at decision points for the rejected class representing unsuccessful projects. Additionally, Shell's experts consider the project's results of high value to the decision-making process within open innovation. Based on the model's results, the lacking insights in the decision-making of the open innovation are supported by the solution design and are particularly reliable as both the performance and validation are good.

Finally, the abovementioned partial conclusions contribute to the answer to which extent the problem is solved. The problem statement of this study for the business case was: *the current open innovation process at Shell does not utilize historical data to gain insights into decision-making capability*. The decision-making within the open innovation process at Shell can be supported by analysing historical data to improve decision-making capability by the combined ML and DMN approach for knowledge-intensive decision-making within open innovation. Here, the DT as ML model identified decision logic at decision points via decision-mining and DMN represented the mined decision logic for visualization and interpretation purposes. The decision-making capability of the process is improved as the new approach improves process transparency, resources can be allocated to the projects to be likely successful, and experts are relieved by a filter model that separates the successful and unsuccessful projects based on the project's characteristics. Shell's experts are still required as the ML model, resulting in the decision model, has a supportive nature. Both the study's qualitative and quantitative evaluation methods confirmed the reliable and powerful insights that this approach can achieve in the specific context at Shell. The study showed how performing currently is in a critical open innovation process at Shell, what its opportunities and limitations are, and how this can be both improved and evaluated. Additionally, historical data is utilized that both gained powerful insights and improved decision-making capability within the open innovation process at Shell. Although the study gains powerful insights and the problem is theoretically solved, the study has limitations and recommendations that could improve the study's current work. These suggestions are discussed in section 5.3 representing the last section of this report.

5.2 Research conclusions

This second concluding section describes the research conclusions and goes more in-depth about the meaning of the study's findings. Particularly, these research conclusions explore the study's findings in context to existing literature and focus on lessons learned. This study aimed to identify the first steps toward an approach that improves decision-making capability within knowledge-intensive innovation processes through the case at Shell. Hence, the previous case study conclusions provide a base for this section. The answers to the Shell-specific sub-questions formulated for the specific case are the input for the general main research question central to this study, which was formulated as:

How can the decision-making within the open innovation process be supported by analysing historical data to improve decision-making capability?

First, the best model is based on a dataset extracted from a dataset with significant-good data quality implying negligible missing values. The combination of good-working ML algorithms and good data quality is not unknown in academic literature. Commonly, using raw data to perform analyses most certainly leads to incorrect decision-making due to poor data quality (Mylavarapu et al., 2019). Data quality is the key factor that influences data analyses in multiple ways, but consistency is one of the most important dimensions (Mylavarapu et al., 2019). Therefore, it can be explained that a model extracted from the dataset with good data quality turned out to be the best model. Also, the efficiency of a DT works best on small datasets and this was followed using only 432 projects in the dataset of the best model. By doing so, the DT returned powerful insights efficiently. This confirms the earlier knowledge that it is required to avoid large amounts of data in DTs (Biard et al., 2015). In association with this implication, the good performance of the DT can be also explained. However, the drawback is that a DT might lack accuracy compared to black-box models (Sachan et al., 2020). The ML model returned an 80% accuracy which is both good and realistic but the precision and recall scores identified that this is good for the rejected class but it is indeed lacking for the converted class as no insights are found. As this study did not use other ML models, e.g. black-box models, it is unknown whether the insights for the converted class could be more powerful.

Moreover, using a DT is in earlier research discussed as highly interpretable and understandable (Sachan et al., 2020), and this is confirmed by the validation method used in this study. Experts who do not have a background in the Information Systems field easily understood the ML model and the mined decision logic represented in the DMN model. However, to make the DT understandable and interpretable, thorough data analyses should be done to extract the correct, required, and as easy as possible readable data for an algorithm. This selection is done by the data analyst (the author of this study) rather than by the ML algorithm itself. Therefore, it can be discussed that a hybrid methodology is still necessary. The Information Systems field tools, as part of Artificial Intelligence, can treat the data perfectly but human knowledge is still required. The Artificial Intelligence tools are powerful but limited as data should be synthesized first by experts because the tools did not return meaningful insights otherwise. A research scope with corresponding data most certainly has many manifested opportunities and limitations and these should be scoped narrowly to get powerful insights. On the contrary, this pre-analysis returns to earlier research that the general benefit of ML is that it obligates data analysts to structure and analyse their datasets appropriately. ML algorithms do not work properly once datasets are not well prepared, which lacks extracting deep insights that are possibly hidden in data (Mylavarapu et al., 2019).

Finally, the answer to the main research question is the combined ML and DMN approach for knowledge-intensive decision-making within open innovation. Hence, the decision-making within the open innovation process can be supported by analysing historical data to improve decision-making capability using this combined ML and DMN approach. Here, the DT as ML model identified decision logic at decision points via decision-mining and DMN represented the mined decision logic for visualization and interpretation purposes. Once the current design and its opportunities and limitations are investigated, a model can be created that fits the data, the existing design, and most importantly the problem defined. The most important lesson learned from the explorative case study is that utilizing historical data provides meaningful insights, which in turn ensures improved decision-making capability (referring to more informed decision-making) in organizations. Especially within open innovation processes as these special types of KiPs are dependent on information and inputs from external parties. However, some assumptions should be made that are important lessons learned as discussed in earlier sections. The generalization, section 4.4, focused on the conditions, assumptions, and resources for generalizing the study's results to a broader context.

5.3 Limitations & Recommendations

This third and last concluding section describes the study's limitations, which are input for the study's recommendations. Therefore, both elements are combined in this section because they are related to each other. To give a short overview of the limitations resulting in associated recommendations, Table 15 shortly introduces the limitations and the subsequent recommendations. The limitations and recommendations stated in the table will be discussed in the remainder of this section.

Table 15. Overview of limitations associated with recommendations

Number	Limitation	Recommendation
1	Scope	Study broader scope
2	Savings	Study possible measurements for investigating savings
3	Data points	Collect more data points
4	Capturing context	Easier automate capturing knowledge-intensive data
5	Theoretical model	Execute implementation phase for complete model validation

First, the study's scope is limited to the defined scope described in section 1.1.3. This study focuses on the GC process as the process of an innovative idea entering the system via an online submission form until (i.e. not including) reaching (or not) stage [8]. However, the new approach might also affect the remaining process after stage [8], because decision-making is filtered using a decision model that might lead to other projects in the remaining stages. This is not the only limitation concerning the scope as the study's focus was on submitting open innovation projects via the online submission form. The online submission form is not the only way of reaching Shell's open innovation process, so other incoming projects might be affected as well because that decision-making might become data-supported using the new approach as well. These other types of inputs are not analysed.

Next, although concrete performances of the ML model are analysed and presented, the concrete performances of the process are abstract rather than concrete. This is due to the lack of parameters to measure performance. Therefore, it is difficult to express process performances in terms of resource savings in numbers. There was not enough opportunity within the study's timeframe to deeply dive into creating techniques for measuring these indications. These crucial data points to calculate concrete savings are missing and are a practical limitation of the study. Moreover, besides the well-fitting results of the new approach, the approach is based on a dataset consisting of data from 2020 on. Therefore, the dataset only consists of data from approximately 2,5 years and the decision outcomes are approximately 25% for converted projects compared to approximately 75% for rejected projects. This might be the reason why the approach did not return desired insights in the converted class of projects. This should be further analysed in the future when more data points are available in the *after GC update 2020* dataset especially more converted projects.

Next, although it did not interrupt this study, it is found that some data are not easily accessible. When expanding the scope to more knowledge-intensive data as input for the ML model rather than characteristics of open innovation projects as input attributes, this is a limitation that is not yet analysed. Capturing knowledge-intensive data is hard within the process as the database does not easily provide outcomes of the expert knowledge. Hence, the decision logic level of the expert knowledge model in Figure 12 was impossible to create and visualize as data was not available within the timeframe of the study. The outcomes of the panels, i.e. the approvals and all attached documents, are hard to capture. One should do a lot of manual transactions in the database to capture knowledge-intensive data in parts rather than the whole result. This implies a lack of automation in capturing this knowledge-intensive data and, as a result, the data is very scattered. Finally, the combined approach of ML and DMN to Shell's open innovation process is still a theoretical model, which is not yet testable in practice. Therefore, complete model validation is not possible due to limited resources. The timeframe of the research should have been longer to thoroughly make the adjustments and measure the positive impacts of changes made. Although a validation for the first model's results is done and both lacking insights into decision-making capability and using historical data to derive decision logic at decision points are captured in the study, a validation of real implementation would give more insights into the actual functioning of the new approach. To conclude, although the study results in powerful insights for both the business case and academic context, the study could be further optimized when expanding the current boundaries.

Bibliography

Ali, N., M. Ghazal, T., Ahmed, A., Abbas, S., A. Khan, M., M. Alzoubi, H., Farooq, U., Ahmad, M., & Adnan Khan, M. (2022). Fusion-Based Supply Chain Collaboration Using Machine Learning Techniques. *Intelligent Automation & Soft Computing*, *31*(3), 1671–1687. https://doi.org/10.32604/iasc.2022.019892

Bagherzadeh, M., Markovic, S., & Bogers, M. (2021). Managing Open Innovation: A Project-Level Perspective. *IEEE Transactions on Engineering Management*, *68*(1), 301–316. https://doi.org/10.1109/tem.2019.2949714

Banu, G. S., Dumitrescu, A., Purcărea, A. A., & Isărescu, S. W. (2016). Defining Open Innovation Concept Using Business Process Modeling. *Procedia Technology*, *22*, 1020–1027. https://doi.org/10.1016/j.protcy.2016.01.135

Bazhenova, E., & Weske, M. (2016). Deriving Decision Models from Process Models by Enhanced Decision Mining. *Business Process Management Workshops*, 444–457. https://doi.org/10.1007/978-3-319-42887-1_36

Biard, T., Le Mauff, A., Bigand, M., & Bourey, J. P. (2015). Separation of Decision Modeling from Business Process Modeling Using New “Decision Model and Notation” (DMN) for Automating Operational Decision-Making. *Risks and Resilience of Collaborative Networks*, 489–496. https://doi.org/10.1007/978-3-319-24141-8_45

Boissier, F., Rychkova, I., & Le Grand, B. (2019). Challenges in Knowledge Intensive Process Management. *2019 IEEE 23rd International Enterprise Distributed Object Computing Workshop (EDOCW)*. https://doi.org/10.1109/edocw.2019.00023

Bramer, M. (2007). *Principles of Data Mining*. Springer Publishing.

Calvanese, D., Dumas, M., Laurson, L., Maggi, F. M., Montali, M., & Teinemaa, I. (2018). Semantics, Analysis and Simplification of DMN Decision Tables. *Information Systems*, *78*, 112–125. https://doi.org/10.1016/j.is.2018.01.010

Clarivate. (2021). *Trusted publisher-independent citation database.* Web of Science Group. https://clarivate.com/webofsciencegroup/solutions/web-of-science/

Czvetkó, T., Kummer, A., Ruppert, T., & Abonyi, J. (2022). Data-driven business process management-based development of Industry 4.0 solutions. *CIRP Journal of Manufacturing Science and Technology*, *36*, 117–132. https://doi.org/10.1016/j.cirpj.2021.12.002

De Carvalho, R. M., Mili, H., Gonzalez-Huerta, J., Boubaker, A., & Leshob, A. (2016). Comparing ConDec to CMMN - Towards a Common Language for Flexible Processes. *Proceedings of the 4th International Conference on Model-Driven Engineering and Software Development*. https://doi.org/10.5220/0005688002330240

Di Ciccio, C., Marrella, A., & Russo, A. (2015). Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches. *Journal on Data Semantics*, *4*(1), 29–57. https://doi.org/10.1007/s13740-014-0038-4

Du Plessis, M. (2007). The role of knowledge management in innovation. *Journal of Knowledge Management*, *11*(4), 20–29. https://doi.org/10.1108/13673270710762684

Elsevier. (2021a). *About ScienceDirect | Premier platform for discovering peer-reviewed scientific, technical and medical information | Elsevier.* https://www.elsevier.com/solutions/sciencedirect

Elsevier. (2021b). *Why choose Scopus - Scopus benefits | Elsevier solutions.* https://www.elsevier.com/solutions/scopus/why-choose-scopus

Etinger, D., Simic, S., & Buljubasic, L. (2019). Automated decision-making with DMN: from decision trees to decision tables. *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. https://doi.org/10.23919/mipro.2019.8756694

Figl, K., Mendling, J., Tokdemir, G., & Vanthienen, J. (2018). What we know and what we do not know about DMN. *Enterprise Modelling and Information Systems Architectures, 13*(2). https://doi.org/10.18417/emisa.13.2

França, J. B., Baião, F. A., & Santoro, F. M. (2012). Towards characterizing Knowledge Intensive Processes. *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. https://doi.org/10.1109/cscwd.2012.6221864

Ghattas, J., Soffer, P., & Peleg, M. (2014). Improving business process decision making based on past experience. *Decision Support Systems*, *59*, 93–107. https://doi.org/10.1016/j.dss.2013.10.009

Herzog, P. (2008). *Open and Closed Innovation: Different Cultures for Different Strategies (Betriebswirtschaftliche Studien in forschungsintensiven Industrien)*. Gabler-Verlag. https://doi.org/10.1007/978-3-8349-8090-8

IBM. (2021). *CRISP-DM Help Overview.* https://www.ibm.com/docs/en/spss-modeler/SaaS?topic=dm-crisp-help-overview

IEEE. (2021). *About IEEE Xplore.* https://ieeexplore.ieee.org/Xplorehelp/overview-of-ieee-xplore/about-ieee-xplore

Isik, Y., Van den Bergh, J., & Mertens, W. (2012). Knowledge Intensive Business Processes: An Exploratory Study. *2012 45th Hawaii International Conference on System Sciences*. https://doi.org/10.1109/hicss.2012.401

Kluza, K., Wiśniewski, P., Jobczyk, K., Ligęza, A., & (Mroczek), A. S. (2017). Comparison of Selected Modeling Notations for Process, Decision and System Modeling. *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems*. https://doi.org/10.15439/2017f454

Lu, J. (2018). A Data-Driven Framework for Business Analytics in the Context of Big Data. *Communications in Computer and Information Science*, 339–351. https://doi.org/10.1007/978-3-030-00063-9_32

Mittal, S., & Sangwan, O. P. (2019). Big Data Analytics using Machine Learning Techniques. *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. https://doi.org/10.1109/confluence.2019.8776614

Mylavarapu, G., Viswanathan, K. A., & Thomas, J. P. (2019). Assessing Context-Aware Data Consistency. *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*. https://doi.org/10.1109/aiccsa47632.2019.9035250

OMG. (2019). *Decision Model and Notation Version 1.*2. OMG Headquarters, 109 Highland Avenue, Needham, MA 02494, USA. https://www.omg.org/spec/DMN/1.2/PDF

Sachan, S., Yang, J. B., Xu, D. L., Benavides, D. E., & Li, Y. (2020). An explainable AI decision-support-system to automate loan underwriting. *Expert Systems with Applications*, *144*, 113100. https://doi.org/10.1016/j.eswa.2019.113100

Saura, J. R., Palacios-Marqués, D., & Ribeiro-Soriano, D. (2022). Exploring the boundaries of open innovation: Evidence from social media mining. *Technovation*, 102447. https://doi.org/10.1016/j.technovation.2021.102447

Scheibel, B. & Rinderle-Ma, S. (2021). Comparing decision mining approaches with regard to the meaningfulness of their results. *ArXiv, abs/2109.07335*.

Shell. (2017). *Innovation and R&D - Shell Investors' Handbook 2012–2016*. https://reports.shell.com/investors-handbook/2016/projects-technology/innovation-and-r-d.html

Shell. (2022a). *Our values*. https://www.shell.com/about-us/our-values.html

Shell. (2022b). *What we do*. https://www.shell.com/about-us/what-we-do.html

Shell. (2022c). *Who we are*. https://www.shell.com/about-us/who-we-are.html

Shell. (2022d). *Shell gamechanger*. https://www.shell.com/energy-and-innovation/entrepreneurs-and-start-ups/shell-gamechanger/about.html

Signavio. (2022a). *Model a decision.* Signavio | The Only All-in-One Business Process Software. https://documentation.signavio.com/suite/en-us/Content/process-manager/userguide/dmn-model-decision.htm

Signavio. (2022b). *Business Process Model & Notation - BPMN Introductory Guide*. Signavio | The Only All-in-One Business Process Software. https://www.signavio.com/bpmn-introductory-guide/

Singh, P., & Manure, A. (2020). *Learn TensorFlow 2.0: Implement Machine Learning and Deep Learning Models with Python*. Apress.

Van Aken, J., Berends, H. & Van der Bij, H. (2012). Problem solving in organisations. New York: Cambridge University Press.

Venero, S. K., Reis, J. C. D., Montecchi, L., & Rubira, C. M. F. (2019). Towards a metamodel for supporting decisions in knowledge-intensive processes. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. https://doi.org/10.1145/3297280.3297290

Verma, P., Awasthi, V. K., Sahu, S. K., & Shrivas, A. K. (2022). Coronary Artery Disease Classification Using Deep Neural Network and Ensemble Models Optimized by Particle Swarm Optimization. *International Journal of Applied Metaheuristic Computing*, *13*(1), 1–25. https://doi.org/10.4018/ijamc.292504

Wu, Y. C., Lin, B. W., & Chen, C. J. (2013). How Do Internal Openness and External Openness Affect Innovation Capabilities and Firm Performance? *IEEE Transactions on Engineering Management*, *60*(4), 704–716. https://doi.org/10.1109/tem.2013.2262050

Wulfen, G. V. (2016). The Innovation Maze: Four Routes to a Successful New Business Case. Laurence King Publishing.

Yin, R. K. (2009). *Case Study Research: Design and Methods*. SAGE Publications, Inc.

Zineb, E. F., Najat, R., & Jaafar, A. (2021). An Intelligent Approach for Data Analysis and Decision Making in Big Data: A Case Study on E-commerce Industry. *International Journal of Advanced Computer Science and Applications, 12*(7). https://doi.org/10.14569/ijacsa.2021.0120783

Appendix

Appendix A Search strategy SLR

This section describes the (re)search strategy for completing the SLR, which consists of the research questions, the sources, the search terms, the selection method and criteria, and the literature found.

Research questions

Throughout this SLR, the following main research question will be addressed:

> *Which (combination of) BPM-modelling approaches or ML techniques can potentially be used for managing knowledge-intensive decision-making within open innovation?*

The following sub-questions are considered to support the main research question of this SLR:

1. *What are the main characteristics of managing knowledge-intensive processes?*
2. *What are the main characteristics of open innovation (in the process industry)?*
3. *Which BPM-modelling approaches for knowledge-intensive decision-making are already used and what are their benefits and limitations?*
4. *Which ML techniques for knowledge-intensive decision-making are already used and what are their benefits and limitations?*

Sources

This section describes the selection of search engines for collecting existing academic and scientific literature to create the literature analysis. The search engines complement each other within the research area regarding articles in journals, e(books), and conference proceedings. Table 16 shows the selected search engines for this SLR, including their descriptions and their variety of offers.

Table 16. Search engines

Search engine (name and webpage)	**Description (including variety of offers)**
IEEE (Institute of Electrical and Electronics Engineers) Xplore Digital Library *https://ieeexplore.ieee.org/Xplore/home.jsp*	Technological digital library offering (IEEE, 2021): > 295 journals and magazines > 6,000 (e)books > 4,000,000 conference papers
Science Direct *https://www.sciencedirect.com/*	Full-text scientific database offering (Elsevier, 2021a): > 2,650 peer-reviewed journals > 42,000 (e)books
Scopus *https://www.scopus.com/*	High-quality database offering (Elsevier, 2021b): > 26,000 active serial titles > 243,400 books > 17,500,000 open access items
Web of Science *https://www.webofscience.com/*	Scientific citation search platform offering (Clarivate, 2021): > 21,000 journals > 126,000 (e)books > 226,000 conferences

Search terms

This section describes the search terms to find appropriate literature to answer the research question and its sub-questions. The search terms consist of two parts: search terms used as inclusion and search terms used as exclusion. The search terms used as inclusion derive from the earlier described research design and ensure boundaries, resulting in the existing literature that is relevant for the literature analysis. However, the search terms used as exclusion do not derive from the research design and were used to prevent selecting wrong or useless existing literature. A combination of both included and excluded search terms are necessary to answer the sub-questions and, finally, the main research question.

Included search terms

Table 17 shows the search terms included to find appropriate existing literature, including to which sub-question it relates. A more detailed explanation of the search terms (and their variants) follow in the remainder of this section. It is important to note that all search terms can expand into plural, as indicated by brackets in the table.

Table 17. Included search terms

Relevant to sub-question	Key search term	Other options search term
1	Knowledge-intensive process(es)	Management of knowledge-intensive process(es)
2	Open innovation	
2	Process industry	
3	Decision modelling	
3	Process modelling	
3	Modelling notations	
4	Machine learning	
4	Big data	
4	Decision support systems	
1,2	Characteristic(s)	Characterizing
3,4	Business data processing	
3,4	Decision(-)making	Decision(-)support
1,2,3,4	Business process(es)	Business

The focus of this SLR is on potential (combination of) BPM-modelling approaches or ML techniques for managing knowledge-intensive decision-making within open innovation. To be able to manage the interconnection of open innovation and knowledge-intensity, the *characteristics* or *characterizing* of both *open innovation* and *knowledge-intensive process(es)* (or *management of knowledge-intensive process(es)*) are included as search terms. To be more specific, this SLR focuses on open innovation within the *process industry*, so this search term is included. Moreover, available literature should be found about *modelling notations* that are already used for knowledge-intensive decision-making. As one of the fundamental topics of this SLR is *decision-making (or decision-support)*, this search term is included for finding BPM-modelling approaches and ML techniques in current literature suitable for decision-making. Additionally, *decision modelling* is added to focus on modelling notations for decision-making in more detail, and *process modelling* is also included to ensure the modelling of processes. Next, *business data processing* is added as a search term to include papers related to the processing of business data, which might result in interesting approaches and/or techniques for answering sub-question three and four. Sub-question four is about ML techniques so, therefore, the search term *machine learning* is included. In addition, *big data* is added as a search term because ML mostly deals with large amounts of data that can be referred to in current literature as big data. To ensure that these types of methodologies do not deviate too much from the main research topic, the search term *decision support systems* is included to find techniques in current literature appropriate for decision-making. Finally, the search term *business process(es)* or *business* is added for all four sub-questions to ensure possible practical implementations that relate to organizations.

Excluded search terms

Table 18 shows the search terms that are excluded to find appropriate existing literature, including a clarification to which sub-question it relates (if it turns out that it is necessary during the search). In the remainder of this section, all search terms (and their variants) will be further explained.

Table 18. Excluded search terms

Relevant to sub-question	Key search term	Other options search term
1,2,3,4	(NOT) Hardware	(NOT) Hardware engineering
1,2,3,4	(NOT) Software	(NOT) Software engineering
1,2,3,4	(NOT) Service	
1,2,3,4	(NOT) Finance	(NOT) Cost

As indicated in the research area, this study focuses on the IS field, which indicates that other disciplines are automatically not relevant. Therefore, this SLR considers approaches and techniques that have potential use within open innovation considering decision-support. Hence, *software (engineering*) and *hardware (engineering)* are excluded from the literature search. Moreover, disciplines within organizations such as *finance* (or *cost*) are not relevant for identifying suitable BPM-modelling approaches or ML techniques within the field of IS. Lastly, IS are about connecting people and technology, which focuses on bringing data and processes together. This results in an exclusion of *service* to prevent papers that focuses on delivering services to customers rather than focusing on in-house data and processes.

Selection method and criteria

As a sequel to earlier described sections, the selection method and criteria section ensure the final boundaries for collecting existing literature. First, the selection method describes the method that is selected for choosing appropriate literature to reduce the number of articles. Thereafter, the selection criteria describe the criteria that are set to decide whether literature is included in this SLR or not.

Selection method

In order to establish this SLR, multiple selection steps should be completed subsequently. This SLR uses a variant of the selection method described in the paper of Mitton, Adair, McKenzie, Patten & Perry (2007). The literature selection approach contains the following three steps:

1. Collection of all (unique) hits from search terms in chosen search engines. This first, initial step results in the so-called *long list*.
2. Selection of collected hits by scanning their titles, abstracts, and conclusions based on selection criteria, which are described in the next section. This is an intermediate step and the resulting list is called the *middle list*.
3. A further selection of collected hits by reading their content based on choosing which hits are most value-added and can be used for capturing content as a possible contribution to answering the research question(s). This resulting last list is called the *short list* and will be used for the literature analysis.

Selection criteria

The selection criteria that are set for this SLR, are divided into two categories: accessibility criteria and content criteria. Content criteria are set to fully match the content of the literature analysis, however, without access, it is impossible to use an article from a journal, (e)book or conference proceeding. Therefore, accessibility criteria are set to check whether an article can be practically used or not. Table 19 shows the inclusion criteria that apply to the list of articles for both categories, which will be further explained in the remainder of this section.

Table 19. Selection criteria

Accessibility criteria	**Content criteria**
Full-text accessibility	Support (parts of) the main topics of the sub-questions
Written in English (or Dutch)	Practice-oriented
	Not older than the 21st century (regarding solutions or approaches)

To guarantee accessibility and usability of the articles, *full-text accessibility* and *written in English (or Dutch)* are chosen as criteria. First, full-text accessibility is required for the last step of the selection method where a final selection is based on a paper's full content. Second, English (or Dutch) language is required to prevent language problems. To be able to select articles with relevant content, an article should *support (parts of) the main topics of the sub-questions.* Next, an article should include a *practice-oriented* approach to ensure possible practical implementations. Furthermore, articles should *not be older than the 21st century (regarding solutions or approaches).* It is important to note that this only holds for solutions or approaches, because the problem and its area are not something from the 21st century and, therefore, articles of an older nature might also be interesting to select.

Literature found

This section elaborates on the literature found using the search strategy described in the previous sections. The search strategy is a variant of the selection method described in the paper of Mitton, Adair, McKenzie, Patten & Perry (2007) and results in a long, middle, and short list. Figure 28 gives an overview of the method and the hits per stage.

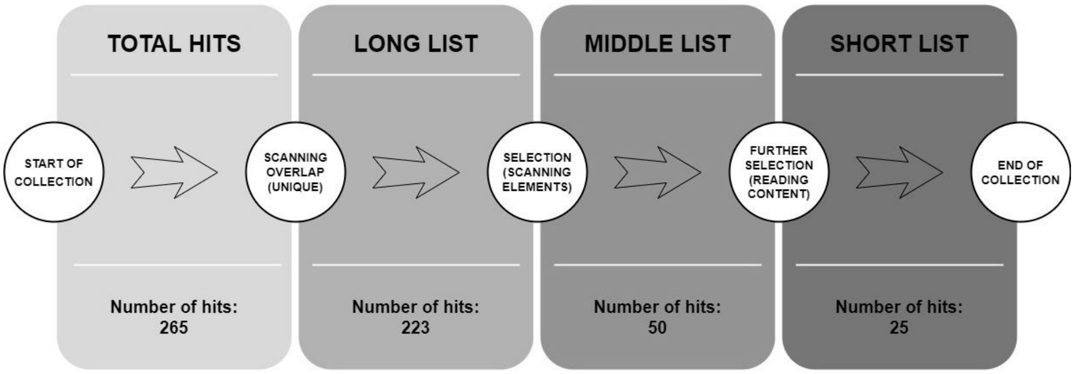


Figure 28. Overview proceeding results all sub-questions

Furthermore, Figure 29 gives an overview of the method and the hits per stage, but this time divided per sub-question (as indicated by the '+' for sub-question 1-4, respectively). It is important to note that after filtering on unique hits, for example, a paper selected at sub-question two does not only hold for this sub-question as this paper might be relevant for other sub-questions as well.

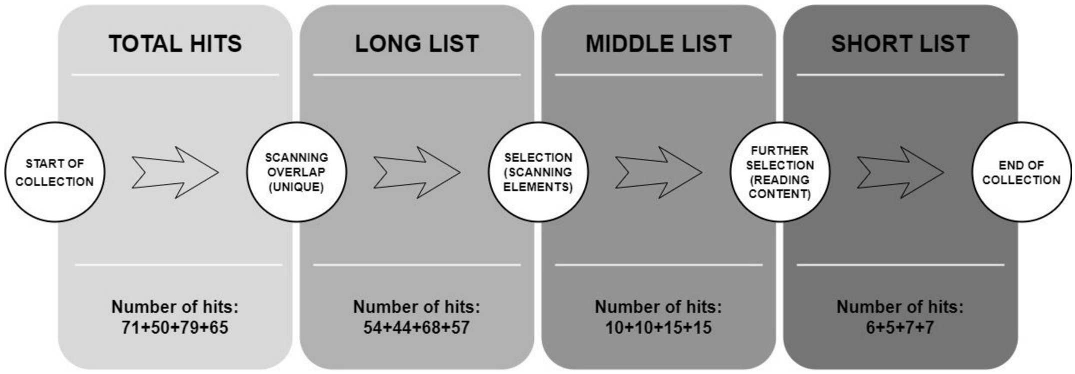


Figure 29. Overview proceeding results per sub-question

Specifically, Tables 20-23 show details of the literature found per search engine per sub-question. The tables give an overview of the following characteristics of the results per sub-question: specific search terms used, the timeframe of publication years, fields used in the search engine, and the number of hits per search engine. Additionally to the earlier described search strategy, more papers are selected based on reference snowballing. These papers are found additionally using the references and citations of the papers in the short lists of sub-questions 1-4 rather than the search strategy beforementioned.

Table 20. Detailed information literature found (sub-question 1)

Search engines / Characteristics	Scopus	IEEE Xplore Digital Library	Science Direct	Web of Science
Search terms	(management of knowledge-intensive processes) AND (characteristics) AND (business processes) AND NOT (software OR hardware OR service)	(management of knowledge-intensive processes) AND (characteristics OR characterizing) AND (business processes) NOT (software OR hardware OR service)	(management of knowledge-intensive processes) AND (characteristics) AND (business processes)	(management of knowledge-intensive processes) AND (characteristics) AND (business processes) NOT (software OR hardware OR service)
Publication years	1989 – 2021	2007 – 2019	2006 – 2021	2002 – 2021
Fields in advanced search	Article title, abstract, keywords	Metadata	Title, abstract or author-specified keywords	Topic
Number of hits	34	9	5	23

Table 21. Detailed information literature found (sub-question 2)

Search engines / Characteristics	Scopus	IEEE Xplore Digital Library	Science Direct	Web of Science
Search terms	(open-innovation) AND (characteristics) AND (business processes) AND (process industry) AND NOT (software OR hardware OR service)	(process industry) AND (characteristics) AND (open innovation) AND (business processes) NOT (software) NOT (hardware) NOT (service)	(open innovation) AND (characteristics) AND (business processes) AND NOT (service)	(open innovation) AND (characteristics) AND (business processes) AND (process industry) NOT (software OR hardware OR service)
Publication years	2007-2022	1997-2022	2008-2022	1994-2022
Fields in advanced search	Article title, abstract, keywords	Metadata	Title, abstract or author-specified keywords	Topic
Number of hits	10	7	6	27

Table 22. Detailed information literature found (sub-question 3)

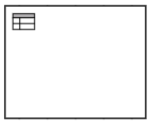
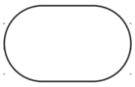
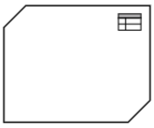
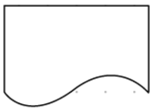
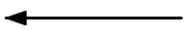
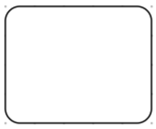
Search engines / Characteristics	Scopus	IEEE Xplore Digital Library	Science Direct	Web of Science
Search terms	(Business process) AND (business data processing) AND (decision making) AND (decision modelling) AND (process modelling) AND (modelling notations)	(Business process) AND (business data processing) AND (decision making) AND (decision modelling) AND (process modelling) AND (modelling notations)	(Business process) AND (business data processing) AND (decision making) AND (decision modelling) AND (process modelling) AND (modelling notations)	(Business process) AND (business data processing) AND (decision making) AND (decision modelling) AND (process modelling) AND (modelling notations) NOT (software OR hardware OR service OR cost)
Publication years	2004-2022	2008-2021	2014-2022	2009-2021
Fields in advanced search	Article title, abstract, keywords	Metadata	Title, abstract or author-specified keywords	Topic
Number of hits	18	28	8	25

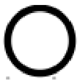
Table 23. Detailed information literature found (sub-question 4)

Search engines / Characteristics	Scopus	IEEE Xplore Digital Library	Science Direct	Web of Science
Search terms	(machine learning) AND (decision making) AND (business) AND (business data processing) AND (decision support systems) AND (big data)	(machine learning) AND (decision making) AND (business process) AND (business data processing) AND (decision support systems) AND (big data)	(machine learning) AND (decision making) AND (business process) AND (business data processing) AND (decision support systems)	(machine learning) AND (decision making) AND (business process) AND (business data processing) AND (decision support systems) AND (big data)
Publication years	2012-2021	2015-2022	2001-2022	2017-2022
Fields in advanced search	Article title, abstract, keywords	Metadata	Title, abstract or author-specified keywords	Topic
Number of hits	14	19	10	22

Appendix B Explanation objects BPMN and DMN

Table 24. Explanation objects BPMN and DMN

Pictogram	Diagram*	Name of object	Description of object**
	DMN	Decision or Sub-decision	Element that determines results based on input data and decision logic, which provides a table for the underlying decision logic.
	DMN	Input Data	Element that provides information for decisions.
	DMN	Business Knowledge Model	Element that adds a function containing business knowledge (e.g., decision table consisting of business rules).
	DMN	Knowledge Source	Element that describes the source of rules for decisions (e.g., guidelines).
	DMN	Information Requirement	Connector between input data or decision element to the specific decision where the information is required.
	DMN	Knowledge Requirement	Connector between business knowledge model to the decision, linking decision logic.
	DMN	Authority Requirement	Connector between any element acting as a source or knowledge to any other element.
	BPMN	Task	Element that represents actions that require to be completed step by step along with the process flow.
	BPMN	Pool/Lane	Elements that define responsibility within a business process, where a pool indicates a boundary to its environment and lanes represent the different roles that execute the process.
	BPMN	Message Flow	Connector for communication that crosses boundaries of a pool.
	BPMN	Sequence Flow	Connector to indicate the order of execution between activities as well as events and gateways.
	BPMN	Exclusive (XOR) Gateway	Element that represents exclusive flow, where the incoming sequence flow is limited by “either/or” the possible outgoing sequence flows based on circumstances.

	BPMN	Parallel (AND) Gateway	Element that represents parallel flow, where the incoming sequence flow multiplies into several outgoing sequence flows.
	BPMN	Start Message Event	Element that indicates the start of the process, which is triggered by receiving an incoming message.
	BPMN	End Message Event	Element that indicates the end of the process, which is followed by sending an outgoing message.
	BPMN	End Event	Element that marks a possible process end, which represents achievement or failure of the business goal of the process.

** Signavio DMN version 1.2 and Signavio BPMN version 2.0*
*** sources: (Signavio, 2022a) & (Signavio, 2022b)*

Appendix C Format project evaluation form

PROJECT EVALUATION FORM

"Improving knowledge-intensive decision-making within open innovation using ML and DMN"

---

1 - How would you describe the degree of well-organized models?

Poor	Fair	Average	Good	Excellent

Explanation (optional):

2 - How would you describe the degree of correspondence to the practice?

Poor	Fair	Average	Good	Excellent

Explanation (optional):

3 - How would you describe the degree of hands-on approach (or do you miss crucial explanations)?

Poor	Fair	Average	Good	Excellent

Explanation (optional):

4 - What do you think of the new model's (redesign) results compared to the old model (current design)?

Poor	Fair	Average	Good	Excellent

Explanation (optional):

5 - What do you think in general of the project's results compared to the (scoped) business value?

Poor	Fair	Average	Good	Excellent

Explanation (optional):

6 - Do you have any other comments or about the project's results?