



CATÓLICA
LISBON
BUSINESS & ECONOMICS

**Business Model Benchmarking:
How a Machine Learning-Based Tool Can
Support Business Model Innovation**

Luca Lenz

Dissertation written under the supervision of Prof. René Bohnsack

Dissertation submitted in partial fulfilment of requirements for the
MSc in International Management at the the Universidade Católica
Portuguesa, 15.06.2020

Abstract

An increasing body of literature has developed around business models and business model innovation in recent years, establishing the concepts' relevance for the competitiveness of a firm. The process of innovating a business model is less researched and still constitutes a complex and challenging task for businesses. Despite the advancements of technology, only few tools have been developed to support said process. Thus, present work focuses on how this process can be facilitated by introducing the concept of benchmarking and complementing it through a machine learning model. A machine learning supported multi-case theory building approach was followed to indicate how benchmarking business models might benefit companies and to arrive at a machine learning model supporting this endeavour. The business models of 306 firms were evaluated to do so. Benchmarking business models can benefit a company by monitoring and learning from other organizations to spark innovation and support idea creation. The machine learning model predicts business model patterns which are used to make processes across companies comparable. The present thesis results in a business model benchmarking tool which supports managers and entrepreneurs alike in their quest of business model innovation.

Keywords: Business Models, Business Model Innovation, Benchmarking, Business Model Patterns, Artificial Intelligence, Machine Learning

Title: Business Model Benchmarking: How a Machine Learning-Based Tool Can Support Business Model Innovation

Author: Luca Lenz

Sumário

Nos últimos anos, um número significativo de estudos foram desenvolvidos em torno de modelos de negócios e na inovação dos mesmos, afirmando a relevância destes conceitos para a competitividade de uma empresa. O processo de inovação de um modelo de negócio não só é pouco estudado, como ainda constitui uma tarefa complexa e desafiadora para as empresas. Apesar dos avanços da tecnologia, apenas alguns meios foram desenvolvidos para dar suporte ao referido processo. Assim, a presente dissertação realça como este processo pode ser facilitado, introduzindo o conceito de benchmarking e completando-o através do modelo '*machine learning model*'. *Multi-case theory building* foi apoiado pelo modelo '*machine learning model*' para indicar como a análise de benchmark dos modelos de negócio podem beneficiar as empresas e para alcançar um *machine learning model* que suporte esse esforço. Nesse sentido, os modelos de negócio de 306 empresas foram avaliados. A aprendizagem e monitorização de outras organizações pode beneficiar o modelo de negócio de uma empresa, estimulando a inovação e a criação de novas ideias. *Machine learning model* é um modelo que prevê os padrões de modelos de negócio que são utilizados para tornar os processos entre empresas comparáveis. A presente dissertação desenvolveu um método de análise de benchmark de modelos de negócio que apoia gerentes e empreendedores na pesquisa pela inovação de modelos de negócio.

Palavras-chave: Palavras-chave: Modelos de Negócios, Inovação de Modelos de Negócios, Benchmarking, Padrões de Modelos de Negócios, Inteligência Artificial, Machine Learning

Título: Benchmarking de modelos de negócios: como um método baseado no *machine learning model* pode apoiar a inovação de modelos de negócios.

Autor: Luca Lenz

Acknowledgements

After nearly two years of studying at Católica Lisbon School of Business and Economics, the conclusion of this work marks another milestone of my life. Starting in September 2018, I could have never imagined what challenges lay ahead, how much I would learn, and how many great friends I was about to make. To everyone who was part of this journey, thank you.

This intense time created new and fastened old friendships which I am thankful of. To mention just a few, I want to wholeheartedly thank my friends and proofreaders, Lea and Vivien, who not only supported me throughout my studies, but in every aspect of my life. No matter how far the distance was. Matthias, for being the best friend in the most difficult times. And of course, my girlfriend Claudia. The person who stood by and motivated, but also intellectually challenged me during all this time. Without you, this time would have never been so bright.

Moreover, I want to thank my supervisor René Bohnsack for his guidance and valuable input which contributed to the high quality of this work. Additionally, to all contributors of his smart city innovation lab and participants of the bi-weekly calls. A special thanks hereby, to the ML expert, our collaboration and discussions have majorly contributed to my comprehension of this challenging topic.

Lastly, I want to thank my family. Without you, this journey would not have even started. I want to thank my brother, who has the wonderful gift of simultaneously cheering one up and relaxing one just with his sheer presence. Most importantly, I want to thank my mother and father. There are simply no words to express how much I owe and thank you. For your unwavering belief in us and your relentless support of whatever idea comes to our mind. For all of this, I am eternally grateful.

Table of Contents

Abstract	1
Sumário	2
Acknowledgements.....	3
Table of Contents.....	4
List of Figures	6
List of Tables.....	7
List of Abbreviations.....	8
Introduction	9
1 Literature Review	11
1.1 Business Models	11
1.2 Business Model Dynamics in the Context of Competition	12
1.2.1 Competition and Business Model Innovation	12
1.2.2 Impediments to Business Model Innovation	14
1.2.3 Business Model Benchmarking.....	16
1.2.4 Business Model Patterns for Business Model Innovation.....	18
1.3 Artificial Intelligence and Machine Learning.....	19
1.3.1 Definition of Concepts	19
1.3.2 Artificial Intelligence in Business and Implications for Competition.....	20
2 Methodology.....	23
2.1 Exploratory Data Analysis.....	24
2.2 Multi-Case Theory Building.....	30
2.3 Machine Learning Support for Theory Building	32
2.4 Benchmarking Processes	34
3 Findings	37
3.1 Machine Learning Results	37
3.2 Multi-Case Theory Building Results and Visualization.....	39
4 Discussion	46

Table of Contents	5
4.1 Managerial Implications	46
4.2 Theoretical Implications	47
4.3 Limitations and Future Research	48
Conclusion.....	50
References	51
Appendix I: Business Model Pattern Description	62
Appendix II: F1-Scores Original ML Model.....	70
Appendix III: Detailed Elaboration on EDA	73
Appendix IV: Results of Upstart Allocation	74
Appendix V: F1-Scores Final ML Model.....	78
Appendix VI: Algorithm Training and Testing Data	80
Appendix VII: Customer Journey	81

List of Figures

Figure 1: Business Model Innovation Process	14
Figure 2: Forms of Benchmarking	17
Figure 3: Usage and Benefits of Artificial Intelligence in Business	22
Figure 4: Count of Predicted BMPs by Original ML Model	25
Figure 5: Top Ten Most Predicted BMPs by Original ML Model.....	25
Figure 6: Accuracy Evaluation of Machine Learning Models	26
Figure 7: Relationship Between Number of BMPs and Market Value	27
Figure 8: Proportional Distribution of BMPs Across Industries	29
Figure 9: Distribution of Industries Within OD, S&P 500 and Upstarts Datasets.....	31
Figure 10: Original Machine Learning Model	33
Figure 11: Model for Conception of a Benchmarking Project.....	35
Figure 12: Detection of the BMPs Experience Selling (ES) and Leverage Customer Data (LCD)	41
Figure 13: Visualization of the Benchmarking Tool, Example: GM.....	45
Figure 14: Process of Compiling the Cases and Utilization.....	74
Figure 15: Customer Journey of Smart Business Modeling Tool	81
Figure 16: Back-End Processes of Business Modeling Tool	82

List of Tables

Table 1: Split of BMP Allocation.....	32
Table 2: Meta-Components for BMP Categorization	36
Table 3: Machine Learning Experiment Results	37
Table 4: Most Employed BMPs Within Upstarts and OD	39
Table 5: Detection of Revenue Sharing and Two-Sided Market Within Datasets.....	42
Table 6: BMP Description Overview	69
Table 7: F1-Score Results of K-Fold Cross Validation, Original Model.....	72
Table 8: Resulting Allocation of BMPs to Upstarts.....	77
Table 9: F1-Score Results of K-Fold Cross Validation, Improved Model	79

List of Abbreviations

AI	Artificial Intelligence
BM(s)	Business Model(s)
BMI	Business Model Innovation
BMP(s)	Business Model Pattern(s)
EDA	Exploratory Data Analysis
GICS	Global Industry Classification Standard
MCTB	Multi-Case Theory Building
ML	Machine Learning
OD	Original Dataset
VCap	Value Capture
VC	Value Creation
VD	Value Delivery
VP	Value Proposition

Introduction

Around the world, the demise of former business giants such as Blockbuster serve as primary examples for the effects of external shocks on the competitiveness of firms. A common source of such shocks are technological changes (Ansari et al., 2016; Benner, 2010; C. M. Christensen & Rosenbloom, 1995; Hill & Rothaermel, 2003; Sood & Tellis, 2011; Tushman & Anderson, 1986). Amplifying this disruptive force or causing it as well are innovative business models (BMs) (Casadesus-Masanell & Zhu, 2013; Kapoor & Klueter, 2015; Markides, 2006). For instance, when Netflix expanded its online business by introducing a high-tech streaming platform, the video rental company Blockbuster fought back by extending its brick and mortar stores and online rental offering (C. M. Christensen, McDonald, Altman & Palmer, 2018). Whereas Blockbuster rented out DVDs and charged excessive late-fees, Netflix employed a complex but customer-friendly ‘*all-you-can-watch, on-demand, low-price, high-quality, highly convenient*’ subscription model (C. M. Christensen et al., 2015, p. 7). Today, out of the formerly 9000 blockbuster stores the company owned in 2004, only one remains (Olito, 2020).

The previous example illustrates how changes in technology and BMs need to be accounted for by companies. Regarding technological change, there is currently no other theme as imminent as artificial intelligence (AI). Although AI is feared to automate jobs, researchers and practitioners alike are working on concepts of how AI can be embraced to reinforce human capabilities, rather than replacing them (Dellermann et al., 2018; Raisch & Krakowski, 2020; Wilson & Daugherty, 2018).

Despite the establishment of the BM in management literature as an organization’s way of doing business, (Cachon, 2019; Täuscher, 2018) and growing research in how to innovate a BM (Foss & Saebi, 2017), changing or innovating one still seems to be a laborious task. The fact that BMs are built upon internal and external factors (Teece, 2018) and include transactions with various stakeholders (Zott & Amit, 2010) complicates business model innovation (BMI). In a focus group conducted by Frankenberger et al. (2013), the interviewed managers stated their difficulties of monitoring important changes and the lack of a tool to ‘*support idea generation for business models*’ (Frankenberger et al., 2013, p. 9). Although IT-tools are primed to support the BM design process (Del Giudice & Straub, 2011), little has been done to do so (Ebel et al., 2016). Thus, in order to address these deficiencies, present work asks:

How can monitoring firms and learning from them benefit the business model innovation process and how can artificial intelligence support it?

Present thesis is structured along four chapters. In the first chapter, a comprehensive literature review delineates the concepts of BMs and BMI and describes how benchmarking, business model patterns (BMPs) and the use of AI can benefit the BMI process. Chapter two outlines how the research questions are addressed, following a machine learning (ML) supported multi-case theory building approach by Tidhar and Eisenhardt (2020). Subsequently, the third chapter illustrates the findings of this study and introduces a ML supported tool for benchmarking BMPs. Lastly, chapter four discusses the managerial and theoretical implications of this thesis and concludes with its limitations and future research suggestions.

1 Literature Review

The following literature review describes business models in the context of competition and investigates the basis and drivers of business model innovation. Subsequently, benchmarking is introduced to facilitate the BMI process. Lastly, the application of artificial intelligence and its managerial benefits for decision-making are outlined.

1.1 Business Models

In recent years, extensive literature has developed around BMs as a concept in management theory (Massa et al., 2017), disproving Porter's (2001) criticism of the BM as '*an invitation for faulty thinking and self-delusion*' (Porter, 2001, p. 13). Besides the research on BMs, a ubiquitous definition of BMs has not been obtained until today (Massa et al., 2017). The most prevalent definitions describe BMs on a meta-level as an activity system of a firm (Zott et al., 2011). An activity system of a firm refers to transactions with various stakeholders such as suppliers and customers (Amit & Zott, 2001; Zott & Amit, 2010). Such an activity can be understood as any transaction of physical or capital resources that a relevant party contributes to achieve a focal firm's objective (Amit & Zott, 2001; Zott & Amit, 2010). Supporting this view, Teece (2010; 2018) describes the BM as the architecture of a firm's activities to satisfy customer needs. Key element of this definition is the focus on the creation of value for customers. Thereby, a business model focuses on how value is created for them, the value delivery as well as capturing said value (Teece, 2010, 2018).

Synthesizing the literature, it becomes apparent that BMs are utilized as a relatively new unit of analysis to investigate the value creation process of firms (Zott et al., 2011). Most definitions of BMs have in common that they focus on the realization of customer needs, how to address them most efficiently, and how to capture the created value (Cachon, 2019; Gassmann et al., 2014; Teece, 2018; Zott et al., 2011). Since Teece's (2018, p. 41) definition of the BM appropriately harmonizes mentioned theories, present work will refer to a BM as:

'The design or architecture of the value creation, delivery, and capture mechanisms it employs. The essence of a business model is in defining the manner by which the enterprise delivers value to customers, entices customers to pay for value, and converts those payments to profit.'

1.2 Business Model Dynamics in the Context of Competition

To understand why and how companies reconfigure their business model, the following chapters elucidate BMs under the aspect of competition and the corresponding process of business model innovation. It further illustrates the drawbacks of business model innovation and opens a way out by introducing the concept of benchmarking and business model patterns.

1.2.1 Competition and Business Model Innovation

Markets and competition are changing with increasing speed and require companies to respond quickly and decisively to new environments (Hermelo & Vassolo, 2010; Yoo et al., 2012). Thus, superior performance and its antecedent competitive advantage are more short-lived for the majority of firms (Wiggins & Ruefli, 2005). Sustainable competitive advantage has become rare and is rather a concatenation of short-term advantages over time (Wiggins & Ruefli, 2005). Additionally, digitization and technological advancements increase the pace of how markets and products change which requires companies to quickly adapt their activities and innovations (Yoo et al., 2012). This, and the more frequent appearances of new competitors in a market, lead to a more intense competition in general (Hermelo & Vassolo, 2010; McNamara et al., 2003).

Managing BMs stipulates one possible option of competing on the firm-level in the modern era (Casadesus-Masanell & Zhu, 2010). Hereby, BMI became a central theme to competition, describing the “*designed, novel and nontrivial changes to the key elements of a firm’s BM and/or the architecture linking these elements*” (Foss & Saebi, 2017, p. 201). Thus, in order to stay competitive or react to new entrants, companies reconfigure their BM (Abdelkafi et al., 2013; Casadesus-Masanell & Zhu, 2013; Saebi et al., 2017), adopt or develop a new BM (Fjeldstad & Snow, 2018; Martins et al., 2015), or build up a BM portfolio including multiple BMs (Ahuja & Novelli, 2016; Eklund & Kapoor, 2019; Tripsas & Gavetti, 2000).

To innovate a BM successfully, a firm should base it on the resources and capabilities of the firm (Barney, 1991; Teece, 2018; Teece et al., 1997) and adjust it to the opportunities in its external environment (Baden-Fuller & Haefliger, 2013; Baden-Fuller & Morgan, 2010). For instance, Kind et al., (2009) find that media outlets adjust their BMs according to their competition. In the absence of substitute products, media companies are more inclined to choose a subscription model since customers do not have an alternative to turn to and thus are considered less price sensitive. If there are one or more substitute products present, media firms rather

bundle their activities to increase the customer base by offering their product for free and using an advertising BM for value capture (Kind et al., 2009).

Alternatively, firms can also diversify their BMs, i.e. employing various BMs in a BM portfolio (Aversa et al., 2017). When employing a BM portfolio, it is imperative for the BMs to be aligned, meaning to share resources and create synergies amongst them (Sabatier et al., 2010). By doing so, the BM portfolio will support sustaining advantages but also spread business risk by diversifying revenue streams and creating valuable capabilities in the future (Casadesus-Masanell & Tarzijan, 2012; Sabatier et al., 2010). Amazon's success in recent years exemplifies latter process (Aversa et al., 2017). Starting with a simple e-commerce BM of selling books, Amazon invested heavily in IT infrastructure and developed the necessary capabilities to employ digital platforms in recent years. At present, the company leverages these capabilities in different business units which employ various BMs such as Amazon Web Services (Aversa et al., 2017). In general, the successful innovation of a BM has been found to potentially build a basis for competitive advantage (Desyllas & Sako, 2013; Frankenberger & Sauer, 2019; Teece, 2010).

To support adjustments to the external environment, firms also base their activities regarding BMI on diverse learning processes (Andries & Debackere, 2013; Foss & Saebi, 2017; Frankenberger & Sauer, 2019; McDonald & Eisenhardt, 2019). Hereby, firms compare own processes to peers and competitors to ensure an effective BM design (Casadesus-Masanell & Zhu, 2013; McDonald & Eisenhardt, 2019). Explicitly, McDonald and Eisenhardt (2019) found that firms which successfully implemented a BM that increases customer value, integrated processes of BMs as prototypes to imitate firms in nascent markets. The prototypes were tested to establish hands-on learning with the company's customers. In addition, firms also borrow attributes from well-established firms which were already legitimized by their customers (Martins et al., 2015; McDonald & Eisenhardt, 2019). For instance, part of Tesla's positioning includes its vertical integration and focus on high-end, luxurious design and technology (Wang, 2018). Thus, similarly to Apple's activity system and its focus on design and technology, Tesla chose to reveal its products on a prominent stage at their headquarter and sells through privately owned high-end stores in shopping-centers (Martins et al., 2015).

A framework of how to approach the BMI process is provided by Frankenberger et al. (2013), as shown in Figure 1. According to the researchers, the BMI process starts with *Initiation* which encompasses the monitoring and comprehension of the external environment of a firm. Hereby, it is crucial to monitor other players regarding their movements and choices. Additionally, the

root causes of changes in the environment should be understood. The second step, *Ideation*, refers to the creation of ideas about how to innovate a BM. The challenges to conquer in this stage consist of overcoming the current business logic, difficulties in thinking in BMs and the lack of tools to develop new ideas. Activities in the *Integration* phase regard the development of a new BM by considering all processes and stakeholders. Lastly, the *Implementation* phase focuses on an iterative process of incorporating a BM within a company. (Frankenberger et al., 2013)

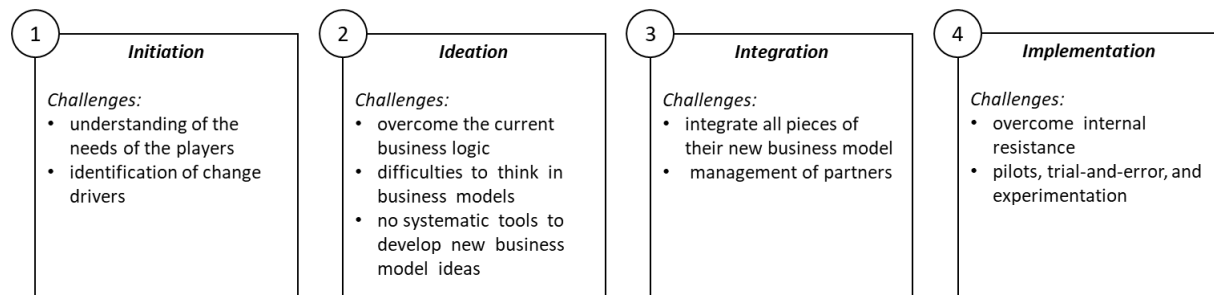


Figure 1: Business Model Innovation Process

Based on (Frankenberger et. al, 2013)

1.2.2 Impediments to Business Model Innovation

While competing through BMs can result in the superior performance of a firm (Desyllas & Sako, 2013), the neglect of BMI can accrue high costs and severe consequences for a firm, impeding the innovation of its BM. These obstacles can be categorized into complexity and associated costs as well as managerial cognition and timing.

As activity systems for value creation, delivery, and capture, BMs are comprised of various, interdependent processes within an organization (Foss & Saebi, 2017). Thus, BMs can be seen as complex systems in accordance to Simon's (1962, p. 468) definition of such systems as various parts which '*interact in a non-simple way*' (Foss & Saebi, 2017). Therefore, changing a component of a BM correlates with a strenuous architectural change which is often accompanied by high costs due to these interdependencies (Foss & Saebi, 2017). Indeed, Eklund and Kapoor (2019) find that innovating a firm's BM comes with significant costs. When incumbents try to adapt to a new entrant with a superior BM, they incur adjustment costs which are caused by implementing new processes and assets while managing the activities of the old value generating BM (Eklund & Kapoor, 2019). Additionally, when incumbents invest in a new model,

their market valuation decreases, resulting in lower liquidity of the incumbent (Eklund & Kapoor, 2019).

Moreover, the emergence of a more value generating BM can result in an inertia of managers caused by their limited cognition (Tripsas & Gavetti, 2000). For example, Tripsas and Gavetti (2000) demonstrate that Polaroid faced the emergence of a new technology which was accompanied with a transformation of the value creating Razor and Blade BM¹. Not accepting the new reality, the managers' inertia prevented the firm from adapting its BM which, in combination with the new technology, eventually resulted in Polaroid's collapse (Tripsas & Gavetti, 2000). Further evidence from the photography industry shows that by changing their BM to a low cost one (No Frills), Kodak was able to challenge the position of the hitherto most successful incumbents in that market, Canon and Sony (C. M. Christensen, 2006; Lucas & Goh, 2009). This transformation established the firm as the biggest player in digital photography at that time (C. M. Christensen, 2006).

Managerial cognition, especially the attention of managers to external factors, can also influence the design and success of BMI. Whereas Frankenberger and Sauer (2019) find that the attention to a certain topic such as a new source of value creation influences the design of a BM, McMullen et al. (2009) argues that managerial attention to competitive intelligence is a determinant of whether a company is able to sustain its performance when facing a competitive threat. Closely linked to managerial attention is the timing of BMI. The literature on competitive responses and dynamic capabilities highlights the importance of making timely decisions when responding to a competitive threat (Barreto, 2010; Bowman & Gatignon, 1995). This was also found to be true for BMI as Kim and Min (2015) denote the timing of BMI as one determinant of its success. However, managerial attention is a limited cognitive resource and thus subject to delusion and faulty thinking (Kahneman et al., 1992; Murphy et al., 1992). This illustrates the need for portraying changes within a company's environment timely and comprehensively.

Whereas monitoring of a firm's environment is critical to commence BMI (Frankenberger et al., 2013), it also determines a BM's design (Frankenberger & Sauer, 2019). Further, being linked to managerial attention and the timing of a decision, monitoring also influences the success of BMI (Frankenberger et al., 2013; Kim & Min, 2015; McMullen et al., 2009).

¹ The sales of a basic, low price product which needs specific consumables to be used. The consumables are highly profitable and contribute majorly to a company's revenue (Gassmann et al., 2014)

1.2.3 Business Model Benchmarking

As previously highlighted, important factors to the design and success of BMI are monitoring a focal firm's environment and learning from other organizations. Benchmarking stipulates an acknowledged practice for businesses and management which allows companies to address shortcomings in their organizational processes by learning from other's best practices (Kunisch, 2017; Watson, 2007).

In general, benchmarking describes the process of comparing one focal object's implicit standing to another reference point (Bamberger & Fiegenbaum, 1996; Festinger, 1954; Zucker, 1987). Its business application denotes the process of comparing the performance and process characteristics between two or more organizations in order to learn how to improve (Watson, 2007). The usage of benchmarking is intended to generate organization-wide core competencies to support achieving a competitive advantage (Bamberger & Fiegenbaum, 1996; Watson, 1993). The process can be partitioned into identifying key areas for improvements and addressing these needs by learning from other organizations' similar operations which are regarded as best practices (Bamberger & Fiegenbaum, 1996; Elnathan & Kim, 1995). Hereby, the organization identifies and studies standards and best practices of other organizations according to which it implements processes and systems to enhance its own productivity and performance (Camp, 1989). Thus, previously described characteristics of the benchmarking process resemble some components of BMI such as understanding and learning from a firm's environment.

Most forms of benchmarking can be categorized as being either based on the nature of the reference point or the process which a firm demands to be improved (Wolfram Cox et al., 1997). Based on said differentiation, Wolfram Cox and colleagues (1997) provide a comprehensive overview of categorizations for benchmarking projects as illustrated in Figure 2. Benchmarking based on the reference point is focused on the organizations a firm compares itself to. The generic *Best in Class* benchmarking relates to the best performing firm, whereas *Industry* benchmarking includes competitors and non-competitors of a given market. *Competitive* benchmarking narrows the reference points down to direct competitors, whereas *Internal* benchmarking only focuses on processes within the company and across different departments. Benchmarking based on processes distinguishes *Global* benchmarking, considering all organizations that employ a process, and *Strategic* benchmarking which focuses on a rather long-term improvement in a process. *Performance* based benchmarking considers the outcome of the

benchmarked process whereas *Functional* benchmarking solely regards a precisely defined process or system. Effectively, these categories can be utilized to understand the theoretical concept of benchmarking but are often combined in practice. (Wolfram Cox et al., 1997)

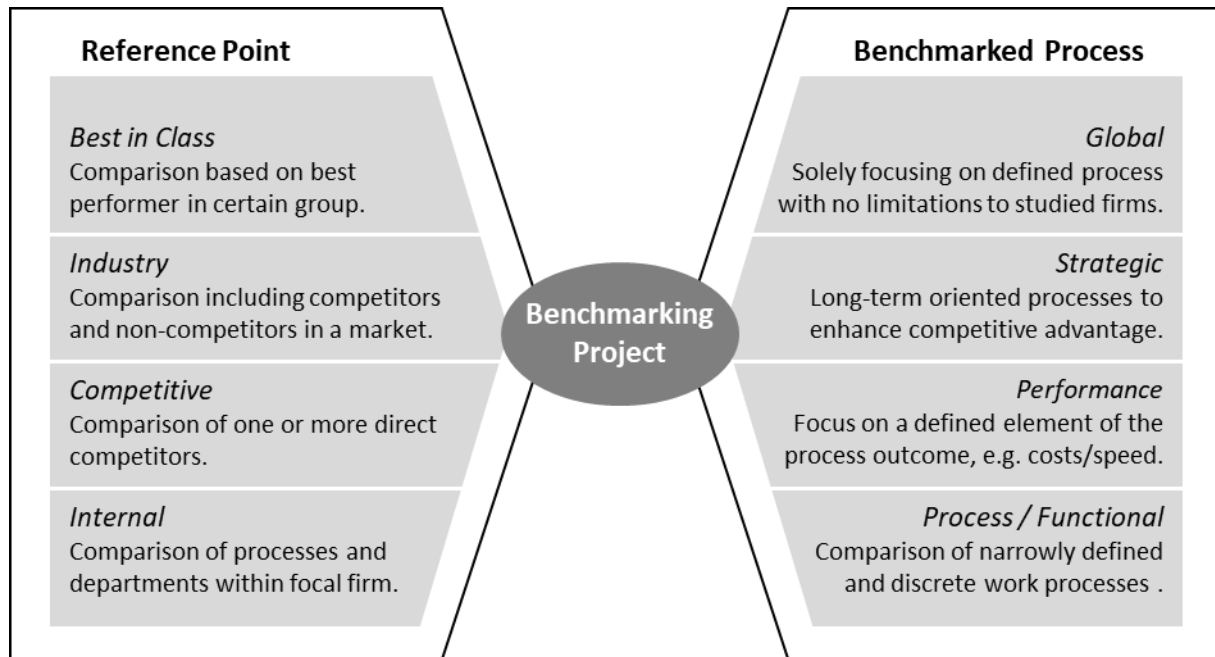


Figure 2: Forms of Benchmarking

Based on Watson (2007) and Wolfram Cox et al. (1997)

Bearing in mind that BMI partly relies on monitoring its external environment and learning from others, there are several interconnections in the literature of how the benchmarking practice can benefit BMI. First, industry benchmarking offers a structured approach to seize and react to external opportunities (Watson, 2007). To facilitate an effective BM design companies study and learn from competitors and other market participants (McDonald & Eisenhardt, 2019). Further, effective BM design includes the consideration of “*the broader business environment*” (Teece, 2010, p. 21). Addressing named characteristics, industry benchmarking contributes to understanding and learning from competitors and non-competitors in a defined milieu. Thereby, it might benefit BMI since the utilization of BMs has been found to be subject to the industry a firm operates in (Saebi et al., 2017).

Second, the interdependency of a firm’s strategy and its activity system of a BM includes various organizational processes to which strategic benchmarking can respond to. Strategic benchmarking could enhance BM design since it focuses on the long-term content of what is benchmarked. Content relates to processes or systems of strategic importance on an organizational level (Watson, 2007). Thus, strategic benchmarking for BMs could benefit the BMI

process by offering a governed approach to monitor and compare how organizations compete through BMs.

Overall, benchmarking could address the need of BMI for monitoring and learning by offering a structured approach. However, a prerequisite of benchmarking are defined processes, contrasting the complex activity systems of business models.

1.2.4 Business Model Patterns for Business Model Innovation

In order to mitigate the complexity of benchmarked processes (Saebi et al., 2017), BMPs offer a possible solution. In general, patterns are defined as solutions to recurring problems (Alexander et al., 1977). Moreover, patterns describe *'the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice'* (Alexander et al., 1977, p. 1216). Integrating this into a BM context, the academic literature conceives BMPs in three pre-eminent ways as:

1. Generalizations of different activities to increase customer value (Amshoff et al., 2015);
2. Solutions to recurring business problems (Abdelkafi et al., 2013);
3. Descriptions of the core of a solution and as such only solving a specific part of a problem (Remane et al., 2017).

Consolidating these conceptions yields in considering BMPs as proven and documented solutions to recurring business problems which can be replicated by others (Amshoff et al., 2015). Further, BMPs might be combined to optimize how businesses generate, deliver, and capture value (Abdelkafi et al., 2013; Osterwalder & Pigneur, 2010). Therefore, a BM of a firm can be replicated by cumulating one or more BMPs (Gassmann et al., 2014).

Academic research on BMPs mostly comprises lists of different applicable patterns (Remane et al., 2017). Hereby, the literature concentrates on a few proven ways to employ BMs, such as Arora et al. (2017) who focus on paid and advertising BMs. Similarly, Rietveld (2018) solely investigates the freemium model. However, this pared representation of value activities is scarcely sufficient to describe the various options of how firms design their activities to generate value (Tidhar & Eisenhardt, 2020).

In contrast, practitioner work often consists of long lists of BMPs which are compiled through studying and observing various companies and conceptualizing their value mechanisms

(Remane et al., 2017; Tidhar & Eisenhardt, 2020). For example, Gassman et al. (2014) composed a BMP list which is able to represent 90% of all BMIs within the last 50 years. Hereby, BMPs are categorized within four dimensions, by defining *who* the target customer is; *what* the value proposition describes; *how* the value chain is structured to produce the respective offering; and *why* it generates value with the selected profit mechanism (Gassmann et al., 2014). However, such long lists – ranging from 55 (Gassmann et al., 2014) up until 182 accumulated BMPs (Remane et al., 2017) – are often based on atheoretical research or coincide with each other (Tidhar & Eisenhardt, 2020). In general, the benefit of thinking in patterns lies in a reduction of complexity and the ability to make different business processes comparable.

1.3 Artificial Intelligence and Machine Learning

The digitalization and the effects of today's data abundance are far-reaching and a major driver of changes (Bharadwaj et al., 2013; Iansiti & Lakhani, 2020; Warner & Wäger, 2019). One theme of technological advancements is artificial intelligence which is utilized not only in a practical context but also in business model research (Dellermann et al., 2018; Szopinski et al., 2019). Following chapters introduce the concepts of artificial intelligence and their applications in business.

1.3.1 Definition of Concepts

The notion of AI originated in computer science (Russell et al., 2010) being defined as “*the study of the design of intelligent agents*” (Poole et al., 1998, p. 1). Thereby, an intelligent agent is a system that acts under perceptual limitations in combination with flexibility towards environmental and goal-related changes considering the appropriateness of the situation, while learning from experiences (Poole et al., 1998). Today, AI is conceptualized to span a broad field with various definitions. To better comprehend the concept, it can be distinguished between *general AI* and *narrow AI* (Broussard, 2018). *General AI* describes software and programs that can think and act fully independently (Raj & Seamans, 2019). However, this form does not exist yet. Therefore, in the present work the connotation of AI is referred to the definition of *narrow AI*, which describes highly elaborated and on algorithm-based procedures supposed to find patterns in datasets and making predictions about the future based on these (Broussard, 2018; Raj & Seamans, 2019).

Machine learning is one subcategory of AI and can be understood as a process of how a computer is trained on data of a specific task in order to improve its performance on said task (Witten

et al., 2017). Within the domain of ML there are three distinguishable learning models (Jordan & Mitchell, 2015):

1. *Supervised Learning* describes a function that maps (x, y) -pairs of an input object x and predicts a desired outcome measure y . The algorithm is trained and tested with a labelled dataset in order to predict a new unlabeled dataset, e.g. an email spam filter categorizes an email as “no spam” or “spam” based on a dataset with ordinary and spam emails (Hastie et al., 2009; Jordan & Mitchell, 2015; Lecun et al., 2015).
2. *Unsupervised Learning* denotes a model that analyses unlabeled data and describes patterns within the dataset assuming general structural properties, but with no outcome measures to guide the process (Hastie et al., 2009; Jordan & Mitchell, 2015).
3. *Reinforcement Learning* describes a model that improves by receiving positive or negative feedback after performing a specific task (Sutton & Barto, 2018).

1.3.2 Artificial Intelligence in Business and Implications for Competition

To comprehend the applications of AI in a business context, it is useful to distinguish between two different domains: *automation* and *augmentation* (Davenport & Kirby, 2016; Wilson & Daugherty, 2018). The more prevalent reference to AI is the process of *automation* which describes largely automated and computer-controlled systems, that can execute repetitive procedures (Raisch & Krakowski, 2020). This eventually results in a previously human-performed task being now undertaken by a machine (Raisch & Krakowski, 2020; Raj & Seamans, 2019). The application of AI as *augmentation* describes the process of a human collaborating closely with a machine and thus making decisions that are supported by AI (Amershi et al., 2014; Raisch & Krakowski, 2020). Latter domain is increasingly applied by various companies, i.e. Microsoft and Deutsche Telekom leverage technology to complement or enhance rather than replace human capabilities (Raisch & Krakowski, 2020).

Hitherto, automation was likely to be prioritized in organizations since it increases operational efficiencies, eventually decreasing costs (Davenport & Kirby, 2016). Prerequisites for automation are easily replicable and highly routinized processes (Raisch & Krakowski, 2020). In addition to process optimization, automation also finds applications in other fields, such as HR (Stephan et al., 2017). At JP Morgan Chase, experienced HR managers worked with an AI-based solution to assess and forecast performance in the field of talent acquisition. Originally designed to supplement managers’ decision-making, the US bank automated the whole process

in order to increase the fairness of the assessment as well as making the process faster and more efficient. (Raisch & Krakowski, 2020; Riley, 2018)

Automation also benefitted strategic functions: Ferreira et al. (2016) used ML to optimize the price setting of an online retailer. Hereby, algorithms were utilized to forecast consumer demand for products which have never been sold before. This was especially useful to optimally price new products and had a positive financial impact on the focal firm (Ferreira et al., 2016).

Whereas automation is aimed to eliminate humans from the equation, augmentation describes a co-evolutionary process of humans and machines (Raisch & Krakowski, 2020). Through the interactions and iterations with humans, augmentation is not as sophisticated in scaling efficiencies when compared to automation (Davenport & Kirby, 2016). However, managers' tasks are often too complex to be modelled in an algorithm which necessitates the inclusion of experts in a certain task (Brynjolfsson & Mitchell, 2017; Raisch & Krakowski, 2020). Although not as widely applied as automation, augmentation is also used in areas such as product innovation (Wilson & Daugherty, 2018). K. Christensen et al. (2017) worked with a text mining algorithm to source comments from external online communities and analyzed them based on users' needs and proposals. The resulting algorithm was then used to filter possible ideas for product innovation and was evaluated by a knowledgeable manager (K. Christensen et al., 2017). Further, augmentation was determined to be beneficial for retailing in areas such as Customer Relationship Management and Supply Chain Management for supporting decisions like inventory planning or make-or-buy decisions (Min, 2010; Shankar, 2018).

In terms of competition, automation improves the speed of standardized processes and ensures greater information-processing, rationality, and consistency (Raisch & Krakowski, 2020). These attributes ultimately result in cost-efficiencies of an organization which are one building block of creating a competitive advantage through AI (Brynjolfsson & McAfee, 2014). In contrast, augmentation necessitates constant human involvement and experimentation (Amershi et al., 2014). Hence, the managers involved in the augmentation process play a crucial role providing the algorithm with their capabilities, intuition, and emotion. The resulting outcome is based on a particular algorithm with human input which is nearly impossible to replicate (Holzinger, 2016). The contribution of augmentation to competition is increased productivity, a higher extent of organizational learning, and fostering innovation (Brynjolfsson & McAfee, 2014; Davenport & Kirby, 2016; Wilson & Daugherty, 2018).

Although the differentiation between automation and augmentation is useful for understanding how AI can impact and benefit companies, in reality the two domains are highly interdependent

and intertwined (Raisch & Krakowski, 2020). For example, considering the aforementioned use of AI in the recruiting process of JP Morgan Chase. Although the algorithm automatically analyses an applicant's profile, it was initially initiated by domain experts (Raisch & Krakowski, 2020). This process is unlikely to stop since the continuous changes in candidate requirements have to be implemented by managers working closely with the algorithm in order to make the model robust (Davenport & Kirby, 2016). Further, combining both paradigms might result in synergies which accrue in additional benefits, i.e. by freeing up resources through automation, augmentation might be enabled which could then be used to identify models for further automation (Raisch & Krakowski, 2020). In fact, by applying both automation and augmentation, organizations are enabled to combine machine rationality with human intuition and cognition which enables organizations for greater learning and adaptability (Calabretta et al., 2017; Raisch & Krakowski, 2020), as depicted in Figure 3.

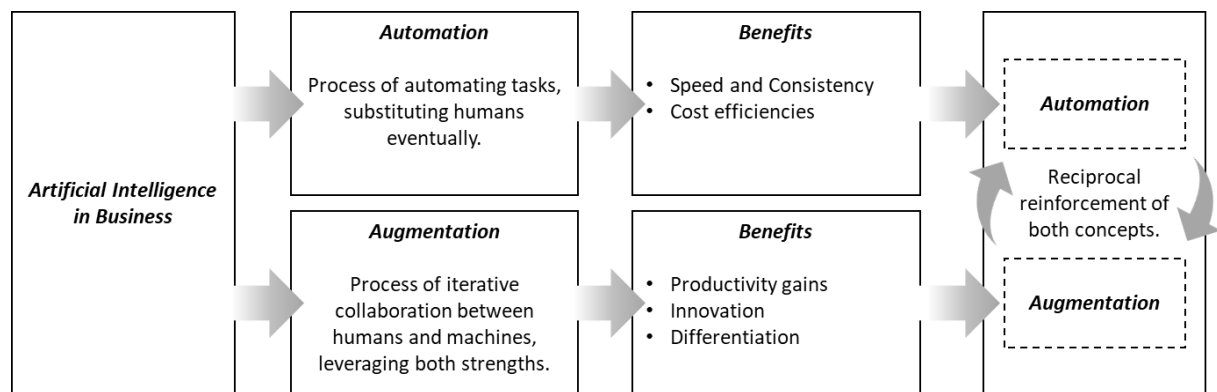


Figure 3: Usage and Benefits of Artificial Intelligence in Business

Considering the benefits of AI to enhance the innovativeness of firms, such as in product innovation (K. Christensen et al., 2017), AI was also found to offer possibilities for BMI. A first advance in this domain undertook Ebel et al. (2016) by developing an IT-tool to support BM design including various stakeholders. Further, ML was utilized by Dellermann et al. (2018) to design a support system for entrepreneurs to evaluate and validate their BMs. Due to the capacity of AI to identify complex patterns, its scalability, and its benefits for innovation, present work utilizes ML to facilitate the benchmarking of BMPs.

2 Methodology

The purpose of present thesis is to illustrate the benefits of benchmarking to BMI and derive a ML supported tool to do so. Therefore, a novel theory building method which was first applied by Tidhar and Eisenhardt (2020) is adapted. The method is comprised of three steps: exploratory data analysis (EDA), multi-case theory building (MCTB), and machine learning. Chosen methodology suits the purpose of present work exceptionally well since the MCTB indicates possible opportunities for BMI which arise from differences in the employment of BMs by companies. The ML model is leveraged to support these findings and can ultimately be utilized for the resulting tool. To visualize the findings, a benchmarking framework is applied.

In order to evaluate BMs and make them comparable, present work utilizes the 55 BMPs which were developed for BMI by Gassmann et al. (2014). This extensive list allows for an accurate replication of BMs and is thereby particularly well suited for the task of benchmarking. An existing algorithm will be analyzed and evaluated in regard to its performance. Said algorithm was created to predict the 55 BMPs to companies based on aggregated company descriptions. Appendix I offers a short explanation of each BMP; more elaborate descriptions can be found in Gassmann et al.'s, (2014) book *'The Business Model Navigator'*.

The exploratory data analysis investigates the original dataset (OD) of 256 firms for which previously mentioned algorithm predicted the BMPs based on their annual reports. To uncover shortcomings and achieve a more accurate understanding of the firms, the dataset was re-evaluated and BMPs allocated by the researcher. The findings of the EDA are elaborated and guide the theoretical sampling of the MCTB. Hereby, an additional 50 firms were evaluated and BMPs assigned based on primary and secondary data. To understand if there are differences in the utilization of BMs, the datasets were contrasted based on their industry affiliation. Further, the aforementioned algorithm was improved through different experiments in order to predict BMPs to the S&P 500 constituents. This dataset is then used to gain further insights for the MCTB. Lastly, a benchmarking framework was utilized for an academically grounded approach to compare and evaluate companies which eventually results in the conception of the BM benchmarking tool.

2.1 Exploratory Data Analysis

The goal of the EDA is to comprehend the given dataset and discover patterns within it (Behrens, 1997). In order to do so, the data is often investigated visually (Tidhar & Eisenhardt, 2020). The process of EDA begins with an investigation through generic categories, breaking them down into sub-categories by defining dimensions along the process until a certain granularity is achieved (Behrens, 1997; Tidhar & Eisenhardt, 2020). Present work will start the EDA by analyzing the dataset which was generated by the previously developed algorithm. Descriptive statistics will be used to understand the data and achieve a first comprehension.

The original algorithm was trained and tested upon a hand-picked dataset of 44 firms which were chosen to represent all 55 BMPs defined by Gassmann et al (2014). After training and testing, the algorithm was then applied to 256 firms to predict BMPs and complement the BMI process. Thus, for starters, the EDA will be based on the dataset of 256 firms to which the algorithm predicted the BMPs. All firms within the dataset are public and based in the United States.

On average, 7.68 (SD = 2.49) BMPs were predicted for each company, resulting in 1967 BMPs in total. The company to which the most BMPs were predicted to was Adobe Inc with 16 BMPs. In contrast, the companies with only two and therefore least allocated BMPs were Honeywell International Inc. and Rollins Inc. The BMPs which were detected the most were *Layer Player* (239), *Solution Provider* (216), and *Customer Loyalty* (204). The 14 BMPs which were not identifiable were: *Add-On*, *Affiliation*, *Aikido*, *Barter*, *Crowdfunding*, *Fractional Ownership*, *Lock-In*, *Pay per Use*, *Pay What You Want*, *Razor and Blade*, *Robin Hood*, *Trash to Cash*, *User Design*, and *White Label*.

In order to get a first overview of the predicted BMPs, the overall count of allocated patterns is plotted in Figure 4. It can be seen that the algorithm has a strong tendency of predicting some BMPs more often than others. This indicates that the algorithm is biased and skewed.

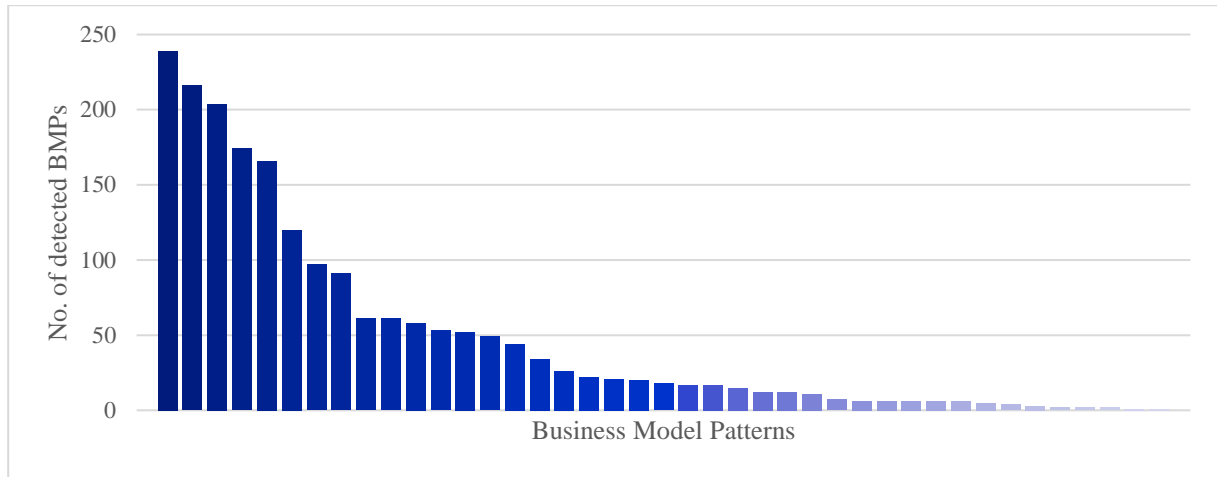


Figure 4: Count of Predicted BMPs by Original ML Model

Zooming in on the ten most predicted BMPs, the dominance of some becomes even more obvious. The three aforementioned BMPs with more than 200 allocations dominate Figure 5. However, the patterns *Target the Poor* and *Franchising* are nearly as dominant with 174 and 166 allocations respectively. Accumulated, the top five most detected BMPs represent more than 51% of all allocated patterns.

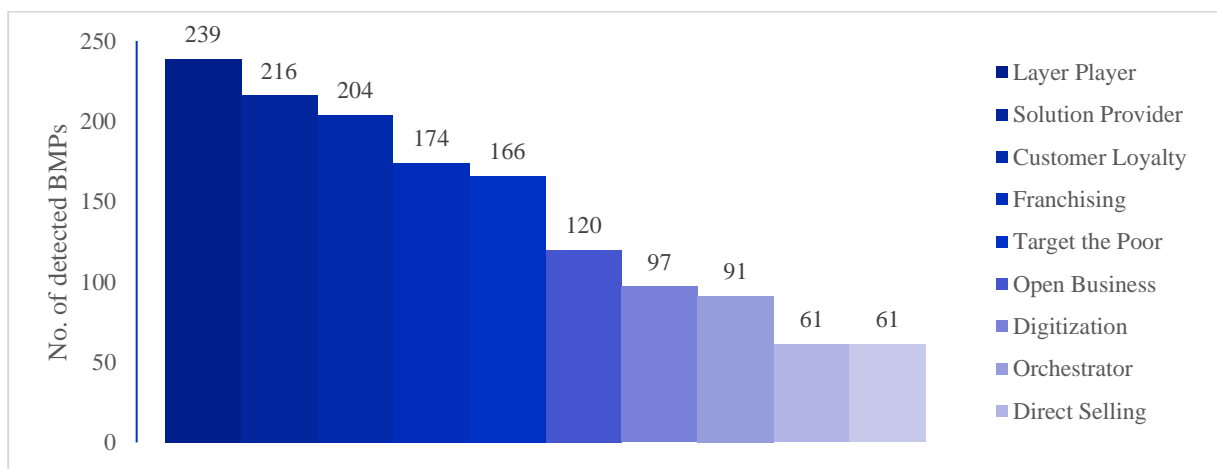


Figure 5: Top Ten Most Predicted BMPs by Original ML Model

Considering the existence of extreme outliers with five BMPs accounting for more than 51% and 14 BMPs not being allocated at all raises doubts on the accuracy of the algorithm.

When evaluating an algorithm which allocates subjects to a given category, its accuracy is used to determine its performance. Thereby, confusion matrices are created as depicted in Figure 6.

		<i>Predicted Class</i>	
		<i>YES</i>	<i>NO</i>
<i>Actual Class</i>	<i>YES</i>	<i>True Positives</i> Correctly predicted positive response.	<i>False Negatives</i> Falsely predicted negative response.
	<i>NO</i>	<i>False Positives</i> Falsely predicted positive response.	<i>True Negatives</i> Correctly predicted negative response.

Figure 6: Accuracy Evaluation of Machine Learning Models

Based on the correctly and falsely predicted responses, the accuracy measures *precision* and *recall* can be calculated as follows (Han et al., 2012):

$$1) \text{ Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$2) \text{ Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Whereas precision stipulates a measure of exactness, recall is a measure for completeness (Han et al., 2012). For instance, precision measures what percentage of positive labelled instances are actually positive, whereas recall measures what percentage of positively labelled instances were labelled as such (Han et al., 2012). Since using one measure isolated can result in an inaccurate model (Han et al., 2012), it is useful to combine them into the F₁-Score:

$$3) F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

The F₁-Score is used since it balances the trade-off between precision and recall and represents a measure to create a balanced classification model (Han et al., 2012). The results of the training and testing of the original ML model are included in Appendix II. These show that some BMPs such as *Digitization* have relatively high scores (F₁ = 0.91) meaning that on average, the prediction of this BMP is in nine out of ten cases correct. However, calculating the average F₁-score for the model reveals an accuracy of F₁ = 0.46 when omitting the BMPs which were not able to be allocated. Thus, the overall accuracy of the algorithm is on average worse than chance for predicting 41 BMPs. This requests a re-evaluation of the companies to gain valuable insights.

In order to balance the inaccuracy and understand the distribution of BMPs across companies and industries better, the whole dataset was re-evaluated by the researchers based on the company's descriptions. In contrast to the allocation of the algorithm, the manual allocation resulted in 1485 total allocated patterns, implying an average allocation of BMPs of 5.80 per company (SD = 2.94). Thus, on average, there were 1.88 less BMPs determined per firm in the sample. The two companies with the most assigned patterns were Microsoft (19) and Alphabet (18). Conversely, the firms with the least allocated patterns were Marathon Oil, and Pioneer Natural Resources Inc, both with one allocated BMP. The most assigned BMPs were *Solution Provider* (164), *Digitization* (108), and *Affiliation* (95). In the dataset of manually allocated BMPs *Crowdfunding*, *Pay What You Want*, *Razor and Blade*, and *Robin Hood* were found to be not employed.

As a first step, the relationship between the usage of total BMPs by a company is plotted against its market value. As it can be seen in Figure 7, there is a slight tendency of having an increased market value by employing more BMPs. However, with Apple, Alphabet and Microsoft employing twelve or more BMPs and being close or above a valuation of one trillion dollars, there are extreme outliers. Further, after investigating the dataset for OLS assumptions, too many observations would have been deleted, making the dataset not suitable for linear regression.

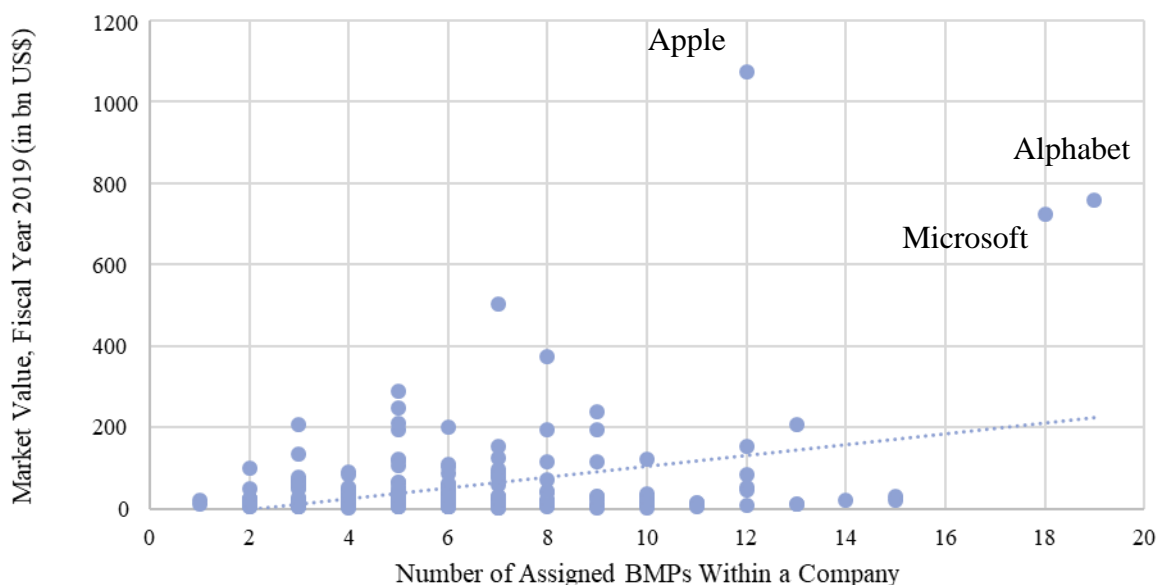


Figure 7: Relationship Between Number of BMPs and Market Value

In order to dissect the dataset, the distribution of BMPs across industries was examined. Research in BMI commonly distinguishes between BMs that are already applied in an industry,

BMs which are new to an industry but already employed in a different industry, and BMs which are entirely new or disruptive to an industry (Foss & Saebi, 2017). This gives rise to the assumption that BMPs vary across industries. Thus, the dataset was segmented according to the Global Industry Classification Standard (GICS) developed by MSCI.

The GICS classifies companies along four categories from eleven industry sectors (included in Figure 9) over industry groups and industries into 158 sub-industries. The GICS is acknowledged for various practical applications by, for example investment banks to classify industries, as well as its solid academic standing (Bhojraj et al., 2003; Kadan et al., 2012). Dissecting for industries suits the approach of present work since companies within an industry can be assumed to share certain characteristics. Thus, a difference in employed BMPs could indicate learning potential for firms regarding BMI. The fact that the two companies with the least allocated BMPs (Marathon Oil, and Pioneer Natural Resources Inc) are both part of the Oil, Gas & Consumable Fuels Industry (GICS: 101020) strengthens this assumption.

As can be seen in Figure 8, which illustrates the distribution of the five most detected BMPs for all industries, there are indeed differences in the utilization of BMPs across industries. One example from the depicted industry sectors will underline this conclusion and be reinforced by a brief evaluation of industry groups. A more comprehensive analysis is included in Appendix III. The BMPs *Direct Selling* and *Affiliation* both describe how firms deliver value. Whereas *Direct Selling* refers to selling to the consumer without any intermediary, *Affiliation* refers to the direct opposite, meaning selling to or buying from an intermediary. The firms within the industry sectors Materials (67%), Information Technology (63%), and Consumer Discretionary (61%) employ *Direct Selling* the most. Contrastingly, the firms within the Financials (65%), Health Care (56%), and Consumer Staples (50%) industry sectors are employing the BMP *Affiliation* the most and it is employed within every industry sector.

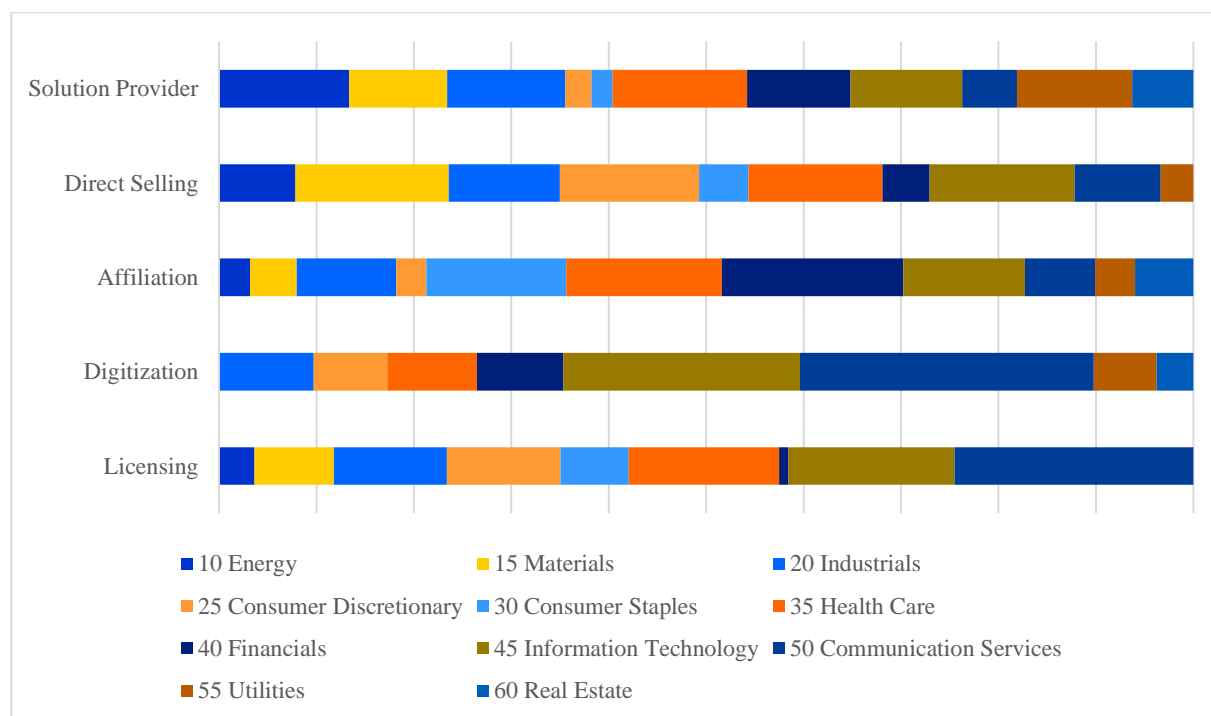


Figure 8: Proportional Distribution of BMPs Across Industries

Increasing granularity by investigating the 24 industry groups according to the GICS adds additional insights. For example, all insurance companies within the sample ($n = 14$) utilize the *Affiliation* BMP, whereas only 21% employed the BMP *Direct Selling*. Further, 85% of the 13 firms within the Pharmaceuticals, Biotechnology and Life Sciences Industry group used *Licensing* as a BMP for value creation. Only the BMP *Make More of It* was used for value capture within this industry group.

It was refrained from increasing granularity even further to the industry or sub-industry levels since the sample size would not allow for valuable insights. Combining the insights of the EDA yields several results to guide MCTB and the remainder of the methodology. First, the algorithm seems heavily biased due to the overrepresentation and extreme outliers. Second, there seems to be a difference in the application of BMPs within different industries. Third, some BMPs seem to be less employed than others even within an industry. Lastly, the accuracy of the algorithm has to be improved in order to facilitate a benchmarking process based on a ML model.

2.2 Multi-Case Theory Building

The purpose of present work's application of MCTB is to draw on cases and investigate differences within the application of BMPs across firms and industries. Such differences could potentially benefit the BMI process of firms. Therefore, present work follows Tidhar and Eisenhardt's (2020) adapted approach of using many "thin"² and mostly cross-sectional cases rather than a few "thick" cases (Eisenhardt et al., 2016). The MCTB approach using many "thin" cases suits present work since the resulting tool will show a snapshot of the currently employed BMPs of considered companies. In contrast, using "thick" cases is usually utilized to understand processes (Tidhar & Eisenhardt, 2020). Additionally, the multiple "thin" cases contribute to the algorithm, serving as a dataset for testing and learning and thus building the groundwork for a possible improvement of the algorithm. Subsequently, 50 firms were chosen using theoretical sampling in order to build the theory and improve the algorithm based on them.

The insights of the EDA and the characteristics of the algorithm were relevant to determine the firms of the new dataset including 50 companies which are thereafter referred to as '*upstarts*'. First, the original dataset including 256 firms does not represent any index although all were based in the US. Since the final algorithm will predict the BMPs to the constituents of the S&P 500 index, 80% of the upstarts were chosen in order to resemble the distribution of industries within the S&P 500, as illustrated in Figure 9. Second, the remaining 20% were chosen to balance the underrepresented BMPs. As illustrated earlier, some BMPs were over- or underrepresented. A significant amount of time was spent on identifying firms that employ underrepresented BMPs to improve the ability of the algorithm to allocate them. Lastly, in order to capture current trends in BM design, the selection of firms was limited to companies that were founded after the year 2000. To guarantee that sufficient and standardized information was available for each firm, and that the firms were relatively successful, the dataset for selection was restricted to firms that went public. This dataset of firms was first assembled by Loughran & Ritter (2004) and continuously updated until March 2020³. For further investigation it was merged with a dataset including variables such as market value or GICS codes

² Content focused cases which are snapshots of companies to a certain point in time and compared to usual cases, rather short and less rich (Tidhar & Eisenhardt, 2020).

³ <https://site.warrington.ufl.edu/ritter/files/2019/05/FoundingDates.pdf>; Accessed: 4th of April, 2020, 10:30 (GMT+1)

obtained from Wharton Research Data Services⁴. This concluded in a sample of 50 firms founded after 2000 which went public until March 2020.

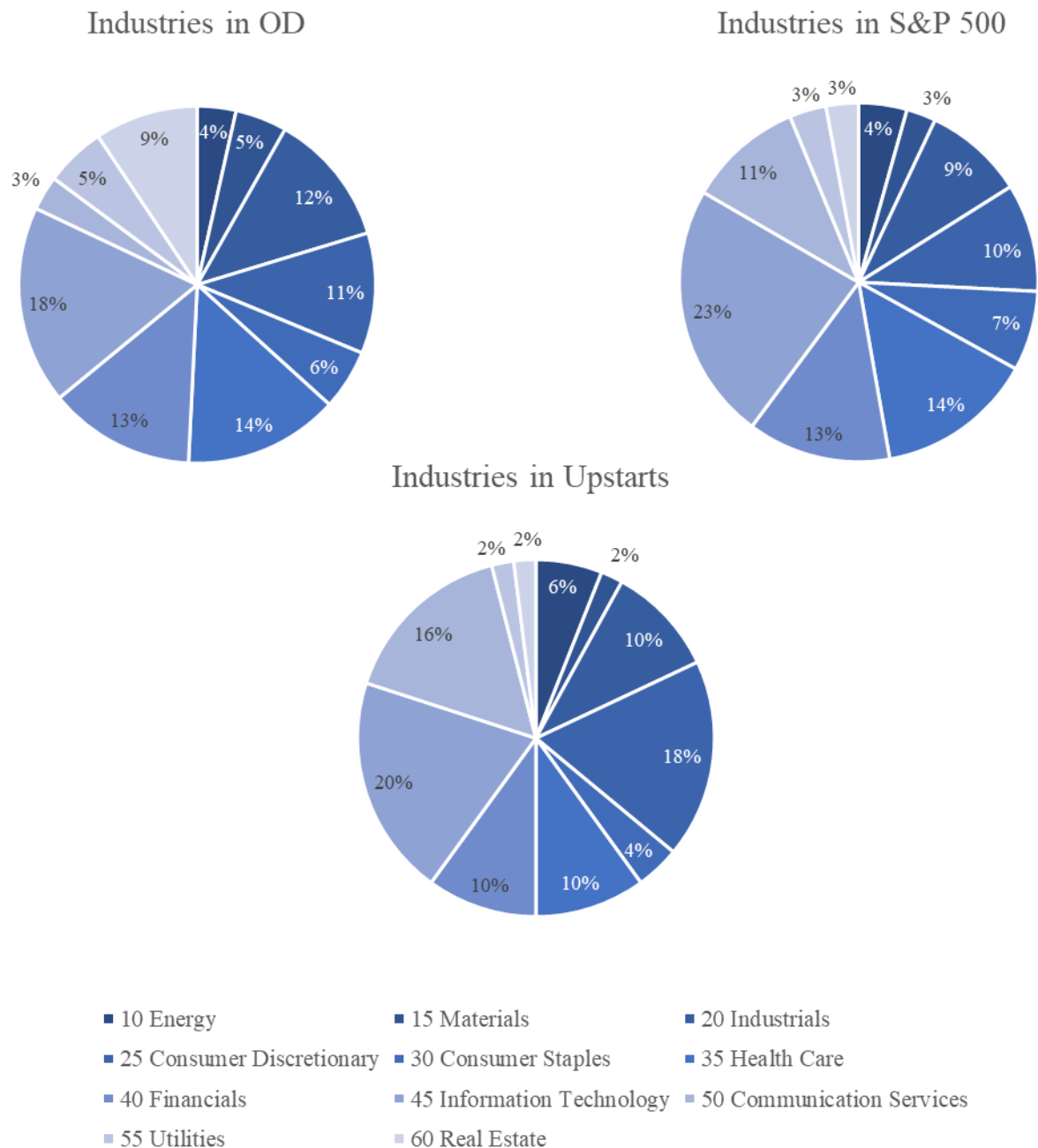


Figure 9: Distribution of Industries Within OD, S&P 500 and Upstarts Datasets

Since the cases are also used for training and testing the algorithm, the case-building was restricted to the specimen of the original ML model. Thus, the cases consist of multiple primary

⁴ <https://wrds-web.wharton.upenn.edu/wrds/>

and secondary sources of company descriptions. Hereby, primary data was evaluated based on annual reports as well as company websites and complemented by the secondary sources of Bloomberg, D&B Hoovers, Crunchbase and Wikipedia. The evaluation distinguished between two categories. *Category 1* denotes BMPs which were clearly assignable. *Category 2* encompasses BMPs which were borderline cases and assigned after a second consideration. The assigned BMPs were then evaluated and affirmed by a BM expert. The resulting allocation is illustrated in Appendix IV.

The within-case analysis was comprised of evaluating the upstarts and assigning the BMPs to them. Moreover, to understand the value activities of firms, a cross-case analysis was conducted. Additionally, to add more scale, the re-evaluated 256 firms of the OD were compared to the upstarts to gain insights in BMP employment. Table 1 provides a brief overview of the overall individual allocation of BMPs to companies to build the dataset.

Respective Dataset	Total of firms within dataset	# Evaluated Firms	Researcher	Validation
Original Dataset	256	50	BM Expert	BM Expert
		206	Researcher	BM Expert
Upstarts Dataset	50	50	Researcher	BM Expert

Table 1: Split of BMP Allocation

2.3 Machine Learning Support for Theory Building

Machine learning is utilized in order to add scale to the findings. Although the application of ML for theory building is a new theme in research, it was deemed to be useful (Tidhar & Eisenhardt, 2020). In this specific case, ML is utilized to find patterns in qualitative data, e.g. company descriptions, in order to arrive at a function that is able to predict BMPs which are used by a given company. Therefore, an existing algorithm is utilized which was previously implemented. The ML model predicts BMPs which are employed by companies based on their description of doing business.

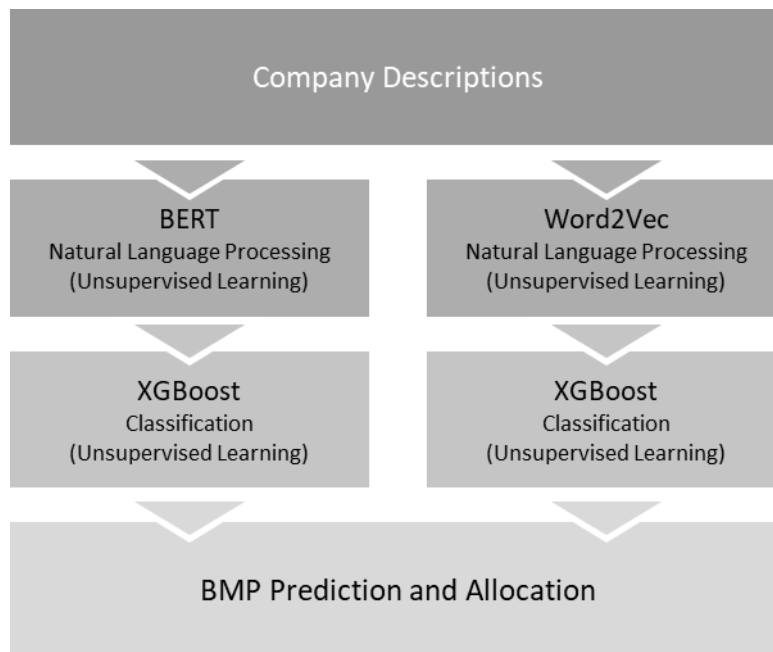


Figure 10: Original Machine Learning Model⁵

As illustrated in Figure 10, the model comprises complementary unsupervised and supervised algorithms facilitating a promising approach to yield higher accuracy in comparison to the usage of a single algorithm. In total, three algorithms are employed within the model. *BERT* (Bidirectional Encoder Representations from Transformers) is an unsupervised algorithm that creates a contextual representation of whole sentences and used for text recognition (Devlin et al., 2018). *Word2vec* is also a text recognition algorithm which creates contextual representations of words within a given sentence (Mikolov, Chen, et al., 2013; Mikolov, Sutskever, et al., 2013). Prior mentioned algorithms are used for Natural Language Processing whose results are then forwarded to the supervised *XGBoost* (eXtreme Gradient Boosting) algorithm. *XGBoost* is a decision tree based supervised learning algorithm and employs gradient boosting which can be used for classification problems (Chen & Guestrin, 2016). Present works' tasks of allocating BMPs to companies based on the company descriptions denotes a language recognition problem as well as a classification problem and thus, these algorithms are appropriate for the given task. As a supervised learning algorithm, *XGBoost* requires training with an appropriate – labelled – training dataset. Meaning in this particular case, company descriptions to which BMPs were assigned to. Additionally, to not confuse the algorithm, the provided training dataset should be forwarded in a standardized format. Thus, the company descriptions which were also utilized

⁵ Original illustration by Astl, (2019). Reshaping Business Modelling With Artificial Intelligence (Unpublished Master's Thesis). UCP Lisbon.

as cases were adapted to the design of the previously aggregated training dataset. The algorithm was then trained and tested in varying iterations using k-fold cross-validation which is a technique that is widely applied in training and testing supervised learning models (Han et al. 2012). Thereby, the dataset was split into 90% of companies used for training the algorithm and 10% for testing. Each iteration yields a mean F_1 -score, the average of those scores denotes the overall accuracy when predicting BMPs.

The original algorithm resulted in an average F_1 -score of 0.46, not being able to allocate 14 out of 55 BMPs. Thus, in order to increase the average F_1 -score, four experiments were conducted in collaboration with a ML expert.

Experiment 1 was built upon the original algorithm using the company descriptions collected by the previous researcher. Increasing the number of observations within a training and testing dataset has often proven to increase the accuracy of prediction. Hence, a second and larger dataset was utilized to further train and test the algorithm. Hereby, the 256 manually allocated companies of the OD were used as labelled training data. However, the algorithm's accuracy did not increase. *Experiment 2* utilized the 44 company descriptions of the previous researcher in combination with the 50 similarly compiled company descriptions of upstarts, compiled by the current researcher. To balance inefficiencies of the previous algorithm, *Experiment 2* did not utilize the previous model but set up a new one solely based on the resulting 94 company descriptions. *Experiment 3* was built upon the resulting algorithm and additionally used a backreferencing technique. Hereby, the researcher referenced where the information of a BMP was included within the compiled company description. Therefore, the company descriptions were able to be subset which reduced the characters for the description of a BMP to be recognized by the algorithm. For the last experiment, solely the dataset of the researcher was used to account for bias in labelling. The iterative process and results are illustrated in the findings.

2.4 Benchmarking Processes

As elaborated in the literature review, the benchmarking process resembles the process of considering external factors and monitoring other companies for BMI. In order to benchmark BMs of companies, present study draws on the proposed framework from Wolfram Cox et al. (1997). It articulates a comprehensive guide for conducting a benchmarking project, illustrated in Figure 11. The model consists of five key components for interorganizational benchmarking and was chosen due to its theoretical grounding and its translatability into a practical context.

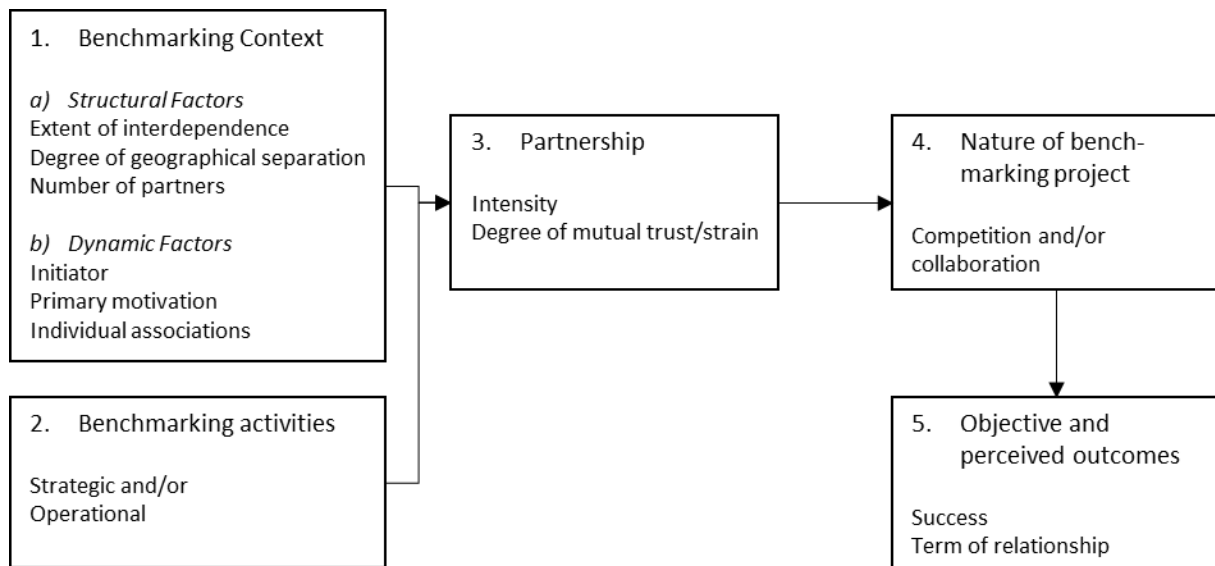


Figure 11: Model for Conception of a Benchmarking Project

Reprinted from Wolfram Cox et al., 1997, p. 298

The *Benchmarking Context* determines the degree of competition or collaboration (Wolfram Cox et al., 1997). Structural factors describe the interdependence between companies and rather static characteristics. Hereby, present work relies on the S&P 500 dataset which limits the companies to the US and the extent of their interdependence to their industries. Dynamic factors describe the ‘*nature and primary motivation*’ (Wolfram Cox et al., 1997, p. 298) of the benchmarking process. Due to previously described relevance of BMPs to firm success, the primary motivation is rather competition based than cooperative.

Benchmarking Activities refer to the regarded processes which are either of strategic or operational nature. BMs are in general related to strategy, often described as its reflection (Casadesus-Masanell & Ricart, 2010). Although the 55 BMPs already reduce complexity, comprehending 55 BMPs is still difficult. Thus, the BMPs were aggregated according to the four meta-components value capture (VCap), value creation (VC), value delivery (VD), and value proposition (VP) used by Günzel & Holm (2013) and Remane et al. (2017), illustrated in Table 2.

The BMPs were categorized according to their descriptions by Gassmann et al. (2014). The categorization was then reaffirmed by the BM expert and is included in Appendix I. The categorization by meta-components was chosen since they allow to gain insights in the value generating processes of firms and facilitate understanding the visualization of the final tool. Additionally, the meta-components allow for insights in how companies structure their process to address their customers.

Meta-Component	Description
Value Proposition	<ul style="list-style-type: none"> • Gives an overall view of a company's bundle of products and services.
Value Creation	<ul style="list-style-type: none"> • Key resources are the assets required to offer and deliver the previously described elements. • Number of key activities performed by key resources. • Some activities are outsourced, and some resources are acquired outside the enterprise.
Value Delivery	<ul style="list-style-type: none"> • An organization serves one or several customer segments. • Value propositions are delivered to customers through communication, distribution, and sales channels. • Customer relationships are established and maintained with each customer segment.
Value Capture	<ul style="list-style-type: none"> • Revenue streams result from value propositions successfully offered to customers. • The business model elements result in the cost structure.

Table 2: Meta-Components for BMP Categorization

Reprinted from Remane et al, 2017, p. 5

The *Benchmarking Partnership* denotes the closeness of relationship between the benchmarked companies. In the present tool, these are determined by the industry relatedness of the GICS codes. Companies within the S&P 500 index are evaluated and visualized according to their respective industry. By doing so, users are enabled to distinguish the relationships between the firms in order to understand the relevance of the firm's BMs. The *Nature of the Benchmarking Project* refers to the extent of competition or collaboration between benchmarked companies. As the tool includes different companies affiliated to a certain industry, the *Nature of the Benchmarking Project* is competitive. This also corresponds to the intents of using the tool for BMI or competitor monitoring. Lastly, the *Objective and Perceived Outcome* can be distinguished by the intent of using the tool, initiating business model innovation and complementing ideation.

3 Findings

The methodology was chosen in order to understand differences within the employment of BMPs by companies and to arrive at an improved algorithm that predicts BMPs to a larger set of companies. Thus, the result section will be split into two parts. First, the resulting algorithm will be introduced and explained. Second, the results of the MCTB will be evaluated and supported by the extended dataset of the algorithm.

3.1 Machine Learning Results

To give leverage to the present work and create a dataset on which the benchmarking tool can eventually draw on, several experiments were conducted to improve the algorithm. The results of the evaluation of the ML experiments are outlined in Table 3.

	Adaptation of Dataset	No. of Identified Patterns (out of 55)	No. of Unidentifiable Patterns (F ₁ -Score = 0)	Mean F ₁ -Score
Original Model	Not applicable	41	14	0.46
Experiment 1	Dataset extension using 10-K forms	45	10	0.43
Experiment 2	Extended dataset with similar descriptions	49	6	0.46
Experiment 3	Backreferencing	12	43	< 0.1
Experiment 4	Controlling for labeling bias	42	13	0.54

Table 3: Machine Learning Experiment Results

The resulting ML model of *Experiment 4* was chosen due to the highest F₁-score. A F₁-score larger than 0.5 implicates that the overall performance of the algorithm is on average better than chance. To get a better understanding of the chosen algorithm, the table illustrating the overall performance of the algorithm is included in Appendix V. It depicts the overall accuracy per F₁-score for each iteration of the k-fold cross validation.

Thirteen BMPs were not able to be identified by the algorithm of the fourth experiment (F_1 -score = 0). Therefore, they were excluded from the ML model. Further, the BMPs *Auction*, *Crowdfunding*, *From Push to Pull*, *Mass Customization*, *Pay What You Want*, *Razor and Blade*, and *Reverse Engineering* had a mean F_1 -score between 0.10 and 0.17. This implies that the model performed below average only predicting one to two out of ten of these BMPs correctly. Further, looking at the number of representations within the upstarts and OD, each of named six BMPs are used by less than 7% of companies. Additionally, within the OD these patterns account for 3.5% of all allocated BMPs and within the upstarts for 2.7%. Due to the low accuracy and utilization of these BMPs, they will be neglected within the ML model. Therefore, the algorithm's final accuracy is a mean F_1 -score of 0.60 after reducing the BMPs to 36.

After the results of the ML were confirmed, the ML expert adapted the model and the prediction of BMPs to the S&P 500 constituents was executed. The final prediction to the S&P 500 companies is included in Appendix VI. Since the BMPs were not manually determined to all the S&P 500 dataset, it is not possible to evaluate the accuracy of the prediction for now.

To draw up the benchmarking tool, the results were merged with additional data on the constituent companies. Further, it was enhanced with the upstart dataset to be able to portray current trends in nascent companies. After the cleaning of the dataset, 467 of the S&P 500 companies were able to be included within the final dataset as well as all upstarts, totaling in 517 firms.

As emerged in the benchmarking methodology, two different dimensions were determined in order to portray the companies and BMPs within the tool: industry relatedness according to the GICS codes and BM meta-components according to Günzel and Holm's (2013) categorization. Due to the varying number of output companies, these are illustrated as rows, whereas the affiliation to respective meta-components is depicted as columns (compare Figure 13).

Since Benchmarking refers to the processes of a Best-In-Class company, illustrated firms are ordered accordingly to their market value. Although the market value cannot be broken down of each BMP's contribution to it, it is useful since it gives an orientation of how the companies compare to each other. The market value was chosen since it accounts for future cash flows and is therefore a forward-looking measure of firm performance (Zott & Amit, 2008). An illustration of the conceived tool is depicted in Figure 13, chapter 3.2.

3.2 Multi-Case Theory Building Results and Visualization

The MCTB was set up to derive insights into the employment of BMPs across and within industries. More explicitly, its purpose is to see whether there are dissimilarities in BMP employment based on the industry relatedness and if yes, whether those insights can be used by firms to benefit BMI. In order to understand possible emergent trends, it investigates the generally most used BMPs within the upstart dataset and then focuses on possible insights within an industry. Due to the sample size of 50 firms, only the industry sectors were investigated.

In total, 450 BMPs were allocated to the 50 upstarts resulting in a mean of 9.00 (SD = 3.16). Thereby, 1.32 more BMPs were assigned to companies compared to the BMP allocation of the original model and 3.20 more compared to the manually assigned patterns of the OD. Roku Inc. was the company with the most assigned BMPs (19), whereas Brigham Minerals and Conifer Holdings were the firms with the least BMPs (3). The most detected BMPs were *Experience Selling* (36 firms), *Leverage Customer Data* (34 firms), and *Digitization* (32 firms). Contrastingly, the patterns *Crowdsourcing*, *Flat Rate*, *Franchising*, *Long Tail*, *Pay What You Want*, *Reverse Innovation*, *User Design*, *White Label*, were not allocated. Notably, there were an additional five other patterns which were assigned only once being: *Crowdfunding*, *Mass Customization*, *Pay per Use*, *Robin Hood*, and *Shop in Shop*.

Most Utilized BMPs by Upstarts	Total	%	Most Utilized BMPs in OD	Total	%
Experience Selling	36	72%	Solution Provider	164	64%
Leverage Customer Data	33	66%	Direct Selling	108	42%
Digitization	32	64%	Affiliation	95	37%
Direct Selling	23	46%	Digitization	89	35%
Affiliation	22	44%	Licensing	76	30%

Table 4: Most Employed BMPs Within Upstarts and OD

As shown in Table 4, the most employed BMPs within the upstart dataset, *Experience Selling*, and *Leverage Customer Data* protrude. Closely following is *Digitization* and the two most

widely used BMPs for value delivery *Direct Selling* and *Affiliation*. The surge of the BMPs *Experience Selling* and *Leverage Customer Data* across all industries is depicted in Figure 12. *Experience Selling* was utilized by 36 out of 50 companies across all industry sectors except Energy and Materials. It describes the process of companies in which they enhance the value of a product or service by offering an experience with it (Gassmann et al., 2014). Although this is often achieved by offering different activities within the distribution channel, such as a coffee corner within a bookstore or book signings, it rather refers to enhancing the value of a product. Thus, the BMP belongs to the meta-component of value proposition. The fact that *Experience Selling* is widely used throughout the upstarts and only by 22% of firms within the OD indicates how companies can benefit from this observation.

Two examples support this line of thought. Sprout Farmer's Market is an American grocery store chain which focuses on selling organically and biologically certified food to comparatively low prices. Simply put, their motto is '*you [don't] have to be wealthy to eat healthy*' (Sprouts Corporate, n.d.). Additionally, it employs the BMP *Experience Selling* by designing the stores to maximize the customer interaction with their employees. Every employee is trained to educate the customers about nutrition and the offered products and ensures that the customers feel welcomed. Thus, Sprouts Farmers Market wants to attract customers not only through their value offerings of natural products, but through the experience of visiting one of their stores. Another example that focuses rather on the product than the sales channel is ETSY. The company operates an online marketplace for selling goods by focusing on the human factor behind each product. Especially, the company emphasizes on the creation of those goods, being hand-crafted by creative people around the world. The company highlights that when you purchase a product on ETSY, you buy something unique. Monitoring the companies' environment could signal that customers are responding positively to an experience focused value proposition and nudge BMI in this direction.

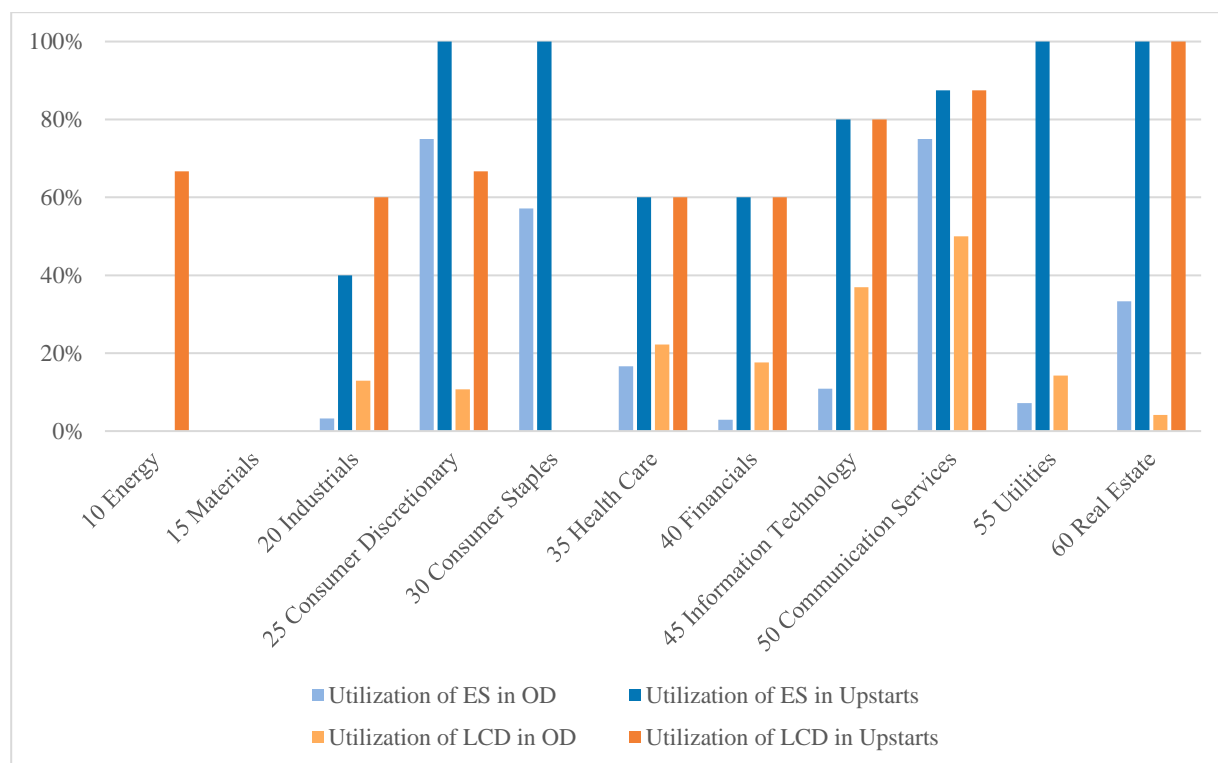


Figure 12: Detection of the BMPs Experience Selling (ES) and Leverage Customer Data (LCD)

Another BMP which promises to be valuable throughout industries is *Leverage Customer Data*. The BMP denotes how firms create value by collecting data about their customers in order to either improve their own value offerings or selling information and additional services to other parties (Gassmann et al., 2014). Since the data is monetized, it belongs to the meta component of value capture. *Leverage Customer Data* is used by 66% of firms within the upstart dataset (33 in total) and just 18% of companies in the OD (45 in total). The BMP seems obvious to the industries of Communication Services and Information Technology. Being digital *per se*, associated companies have a natural abundance of data. To exemplify how firms from other industries might benefit from this BMP, the Pharmaceutical company 10X Genomics serves as an example. The company offers software which compiles data from researchers, analyses it and presents it in an easier to understand manner. Another case in point is the used-car retailer Carvana. By employing a set of algorithms based on the behavioral data of its customers, Carvana optimizes its inventory and acquisition of further cars. The divergence between companies of the OD and upstarts might hint that organizations should start utilizing their customers' data to capture more value.

Moreover, a surge of the BMPs *Two-Sided Market* and *Revenue Sharing* can be observed and gives reason to investigate their utilization in isolation and combined, as illustrated in Table 5:

BMPs	Upstarts (n=50)		OD (n=256)		S&P 500 (n=467)	
	Total	%	Total	%	Total	%
Revenue Sharing	16	32%	5	2%	304	65%
Two-Sided Market	21	42%	26	10%	241	57%
Companies combining both BMPs	11		1		187	

Table 5: Detection of Revenue Sharing and Two-Sided Market Within Datasets

The *Two-Sided Market* BMP refers to a platform which facilitates transactions by connecting the demand and supply-sides (Gassmann et al., 2014). Its processes are based on a company's key activities and belongs therefore to the meta-component of value creation. *Revenue Sharing* refers to the principal of collaboration between firms, such as one company operates a platform on which it invites another company to participate and share the benefits (Gassmann et al., 2014). Both BMPs consider involving multiple parties which indicate synergies. One possible example hereby is Zynga. The company focuses on social gaming by integrating its games in social media platforms. However, it also operates a platform on which Zynga itself, but also other developers publish games. The admittance of other publishers increases the user count which in turn leads to higher revenue for the publisher and more traffic on the platform which then can be monetized through different activities. As seen in Table 5, upstarts are leveraging these synergies with eleven out of 16 companies (69%) that use *Revenue Sharing* also employing the *Two-Sided Market* BMP. Further evidence can be drawn from the S&P 500 dataset. While the BMP *Revenue Sharing* was predicted to 304 BMPs, the BMP *Two-Sided Market* was detected for 241 companies. Of those 304 companies, more than 61% firms (187) also employed *Two-Sided Market*, indicating a beneficial combination of both BMPs.

Investigating industries also illustrates differences in BMPs which can be observed when monitoring companies. The utilization of *Licensing* by upstarts (18%) compared to the companies within the OD (30%) decreased by twelve percentage points. Curiously, when dissecting the data into industry groups, two trends emerge. First, within the Pharmaceuticals, Biotechnology & Life Sciences industry group, the dominant BMP for value capture is *Licensing* with 85% of companies within the OD employing it. Companies in the OD only employ *Make More of It* as

an alternative BMP for value capture (29%). Despite *Licensing* being also employed by companies within the upstarts of the same industry group, 10X Genomics and Quanterix both utilize the BMP *Razor and Blade* in order to create additional revenue through selling consumables for their instruments. On the other hand, companies within the Media & Entertainment industry group which traditionally also use *Licensing* found other ways to capture value such as *Hidden Revenue*, *Revenue Sharing*, and especially *Subscription*. One extreme example is the platform by Workday Inc which explicitly states that it switched from a *Licensing* BMP to a *Subscription*-based one. This finding is also detected by the prediction of the ML model. Within the S&P 500 dataset, the BMP *Licensing* was predicted for only 36% ($n = 25$) of the firms within the Media & Entertainment industry group. In contrast, the BMP *Revenue Sharing* was predicted for 76% and *Subscription* for 68% but *Hidden Revenue* only for 32% of companies within this industry group.

Finally, there are also differences between direct competitors. For example, Tesla was analyzed within the dataset of upstarts and can be compared to General Motors (GM) within the OD, being both automobile manufacturers. The total amount of BMPs employed is similar, with Tesla and GM employing twelve. Further, both companies have a fairly similar distribution of BMPs across the meta-components value proposition, value creation and value delivery with Tesla having four, five and two and GM four, five and three, respectively. The differences lie within the BMPs. Whereas Tesla focuses on selling its products through vertically integrated retail locations, GM utilizes the *Affiliation* BMP of selling through independent dealerships. Further, Tesla sees itself as a vertically integrated technology company. This translates into its usage of BMPs. The company employs the BMPs of *Digitization* and *Leverage Customer Data*, underscoring its focus on technology, whereas none of those were used by GM. *Leverage Customer Data* is also the only BMP for value capture between the two companies and not employed by GM. Thus, Tesla and GM serve as a prime example of how the comparing BMPs of firms might induce stimuli for BMI.

To deepen this analysis and illustrate the final benchmarking tool, Figure 13 portrays an example of the previously described automotive manufacturer GM in comparison with its competitors, the upstart Tesla and incumbent Ford. The BMPs of Ford were predicted by the final algorithm. To get a better overview, BMPs which were employed by all firms are illustrated with white font on blue background. BMPs which were used by more than two are depicted with black font on an orange background. Increasing the scope of the analysis, the upper left-hand graph depicts detected BMPs of companies which are associated to an industry

that includes ‘Automobile’ or ‘Automotive’. Additionally, to illustrate possible trends, companies within the dataset of upstarts and belonging to the same industry sector as GM⁶ were analyzed in the upper right-hand graph. Hereby, the meta-component value proposition dominates in both analyses. In contrast, value capture is in general the least detected meta-component. As illustrated in Figure 13, the company Ford, as the only other company in the ‘Automobile Manufacturers’ sub-industry (GICS: 25102010), heavily diversifies its employed BMPs. Ford outscores its competitors by employing ten more BMPs (GM: 12, Tesla: 12, Ford: 22). However, Tesla and Ford employ both at least one BMP for capturing value, indicating a possibility for GM to add a BMP in said realm.

⁶ Consumer Discretionary

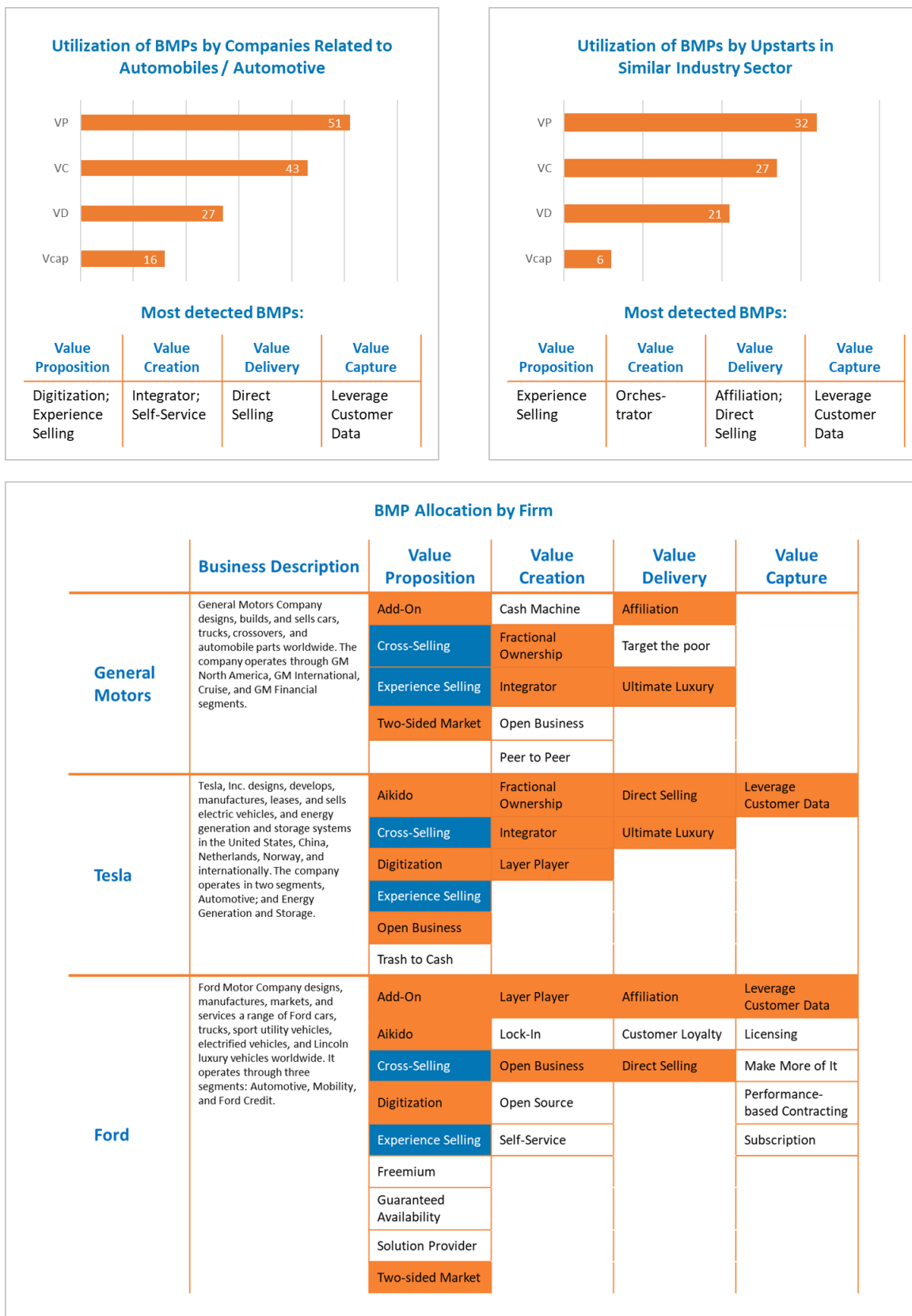


Figure 13: Visualization of the Benchmarking Tool, Example: GM

4 Discussion

Present work's purpose is to understand if and how companies can capitalize on insights from business models of other companies. It introduces the notion of BM benchmarking as a structured approach to monitor organizations and learn from them in order to support BMI. Therefore, company descriptions of 306 firms were evaluated and BMPs assigned to. This resulted in an understanding of how benchmarking can benefit the BMI process and an improvement of the used ML model for predicting BMPs. To elaborate on present work's research question, the managerial and theoretical implications are discussed, concluding with present work's limitations and suggestions for future research.

4.1 Managerial Implications

The academic literature indicates that external stimuli affect BMI and can be used to eventually support managerial decision-making. Hereby, monitoring other companies and learning from their BMs were deemed important (Frankenberger et al., 2013; McDonald & Eisenhardt, 2019). Benchmarking offers a structured approach to facilitate such processes (Watson, 2007). To mitigate the complexity of value generating processes, Gassmann et al.'s (2014) BMPs were introduced. Combining the 55 BMPs allows to replicate a firm's complete BM (Gassmann et al., 2014). Additionally, a framework to design benchmarking projects was utilized to derive a tool that monitors BMs of companies. Further, to respond to the cognitive constraints of managers, a benchmarking tool was introduced based on a ML algorithm which predicts BMPs. The ML support alleviates the laborious task of allocating BMPs.

The MCTB was primarily conducted to understand whether a comparison of BMPs yields meaningful results and can be useful for companies. For instance, it was shown that some BMPs are increasing in popularity across industries (*Experience Selling, Leverage Customer Data*), whereas the usage of others seems to be substituted (*Licensing*). Additionally, a comparison of BMPs on the company level was conducted on the example of GM, Tesla and Ford which indicated a shift in BMPs. These findings were majorly supported by the ML model.

As Bohnsack et al. (2014) point out, incumbents and entrepreneurial firms differ when innovating their BM. Thus, the resulting tool is conceived for two user-groups. First, it can be used to support incumbents in their BMI process. Hereby, the tool addresses the BMI process stages *initiation* and *ideation* (Frankenberger et al., 2013). It supports initiating BMI by monitoring a company's business environment and enables managers to identify change drivers by being

built upon a large and further scalable dataset. To support ideation, the illustration of BMPs of other successful companies and different industries can augment the awareness of managers on different ways of doing business and potentially break the dominant logic. Additionally, the tool is specifically tailored to BMPs which not only mitigates the problem of difficulties thinking in BMs but also gives them an employable tool at hand.

Second, it can be utilized by entrepreneurs. First of all, the tool is able to depict companies within a given industry and can thus identify potential competitors. Second, since startups often test and iterate a BM when trying to discover the most value generating mechanisms (Ghezzi & Cavallo, 2020; Trimi & Berbegal-Mirabent, 2012), the tool can support defining a start-up's positioning by screening a given industry. On the one hand, it enables the user to analyze potential shortcomings of already established companies in their value generating activities. Such insights could then be used to derive possible actions for choosing their business model. On the other hand, it can also show which BMPs have traditionally worked well within an industry.

Finally, the tool will be implemented into a smart business modelling platform. An outline of a user's customer journey and its back-end processes is included in Appendix VII.

4.2 Theoretical Implications

The contributions of this work to BM literature are three-fold. Firstly, it introduces the notion of BM benchmarking. BM benchmarking describes how the comparison of BM-practices within and across industries can facilitate learning and drive BMI. By comparing BMPs across industries, BM benchmarking responds to multiple acclaimed needs, e.g. by Zott and Amit (2008), to further understand how companies utilize and manage their BMs for competition. It also responds to Frankenberger et al. (2013), offering a tool to support the ideation phase of BMI.

Secondly, it investigates the value generating activities of firms. By giving practical insights which of Gassmann et al.'s (2014) 55 BMPs are actually utilized, a practical perspective on reducing lists of BMPs is introduced. Thereby, it offers an alternative, practically oriented approach to Remane et al.'s (2017) taxonomy.

Thirdly, this thesis adds to the emerging literature on how IT-tools can support BM development. For example, Szopinski et al. (2019) state the need for a tool that complements idea generation. Therefore, current work not only utilizes BMPs but also offers insights of how other companies use them to design their BM.

Moreover, present work also contributes to the establishment of a fairly new methodology approach, introduced by Tidhar and Eisenhardt (2020). The application of ML to understand practical cases and support theory proved useful. Especially intriguing is the possibility to draw up datasets and use them for statistical analysis which was otherwise hardly feasible.

4.3 Limitations and Future Research

Present work encompasses several limitations. First, the allocation of BMPs to a company is highly subjective. As Tidhar and Eisenhardt (2020) point out, the BMPs are not mutually exclusive and rather loosely defined, complicating the allocation of BMPs. Moreover, it poses an obstacle for the replicability by other researchers due to a possible misalignment in the comprehension of the BMPs. To mitigate this, the BMPs were cross-validated by a BM expert. Additionally, a list of rules was established in the process, describing borderline cases and illustrating examples for future researches to refer to. However, it does not fully account for the limitation.

Second, the allocation of the BMPs is solely based on the company descriptions and might not reflect the true BM of a company. Since the descriptions of the companies and the corresponding BMPs are the basis for training and testing the algorithm, it is essential that no external knowledge is referred to when allocating the BMPs to a company. Otherwise, it could confuse the algorithm and impair its accuracy and generalizability. Therefore, annual reports of companies were enriched with descriptions from other sources. However, it cannot be guaranteed that the assigned BMPs replicate the complete BM, a company actually employs.

Lastly, the method of benchmarking BMs is based on academic research but was not validated with practitioners. Although the concept responds to various acclaimed needs in the literature, the assumptions were not validated by any users, yet. Additionally, the benchmarking tool is only to be considered to support the BMI process by indicating changes. It does not outperform the accuracy of a human allocating the BMPs manually and only gives indications of BMPs which might enhance the process, not making any causal relationships. This gives rise to future research. Referring to the tool's usage by practitioners, it needs to be investigated how managers and entrepreneurs can benefit from it.

Further research also needs to be conducted in the realm of BMP categorizations. Seeing companies through a BMP-lens has the benefit of making value generating processes comparable across companies. When categorizing and aggregating these BMPs, common trends can be analyzed due to a larger scale. Thereby, the categorization dictates which insights are drawn from

the data. For example, the BMPs could be categorized according to Porter's generic strategies or extent of digitization to increase the comprehension of a firm's competitors and industry.

Finally, the resulting dataset and algorithm open the doors for deeper statistical analysis. Several areas come to mind such as the influence of number of BMPs employed on market value. Further, when the dataset is enriched it could be also interesting to investigate the characteristics of firms which employ certain BMPs. In combination with the market value, such characteristics might support a prediction, whether a BMP is beneficial for a company or not.

Conclusion

Market landscapes around the world are changing at an increasing pace. These developments are partially caused by technological advancements, such as artificial intelligence, which cause lines between industries to blur and increase competition (Iansiti & Lakhani, 2020). Thereby, managers and academia are observing how firms adapt their business models in order to respond to these changes.

Present work responds to those trends by answering the question how benchmarking business models can benefit business model innovation. Since BMI relies on considering external factors of a focal firm, including the monitoring and analysis of other organizations, benchmarking offers a structured approach to facilitate this process and learn from others. By an extensive analysis, using a methodological approach of three steps that build up on each other, it is shown that there are differences how firms design their BMs which could potentially support the BMI of a firm.

Moreover, this thesis presents a tool that is able to screen industries, the companies within them and determine employed BMs of respective firms. The proposed tool utilizes a ML algorithm that allocates BMPs to companies based on their annual reports. Although the accuracy of the ML model was increased by 0.14 to a mean F_1 -Score of 0.60 it is important to understand that the tool does not outperform the accuracy of a human who determines BMs. However, considering the excessive amount of time being spent on understanding the BMPs, evaluating a company and finally assigning the BMPs to the firm manually, the algorithm and resulting tool offer value by augmenting managers in their BMI endeavor.

The conceived tool illustrates how organizations design their BM which is exemplified by three automobile manufacturers. On the one hand, it can ignite the innovation process of a firm, by illustrating movements within a given industry. On the other hand, the tool can support the ideation of managers supporting them to think in BMs and point out different avenues of innovation. Finally, the tool also offers value for entrepreneurs to evaluate certain companies or industries and propose possible BMPs to enhance the value creation, delivery and capture of their firm.

References

- Abdelkafi, N., Makhotin, S., & Posselt, T. (2013). Business model innovations for electric mobility-what can be learned from existing business model patterns? *International Journal of Innovation Management*, 17(1). <https://doi.org/10.1142/S1363919613400033>
- Ahuja, G., & Novelli, E. (2016). Incumbent responses to an entrant with a new business model: Resource co-deployment and resource re-deployment strategies. *Advances in Strategic Management*, 35, 125–153. <https://doi.org/10.1108/S0742-332220160000035006>
- Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A pattern language: towns, buildings, construction*. Oxford University Press.
- Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4), 105–120. <https://doi.org/10.1609/aimag.v35i4.2513>
- Amit, R., & Zott, C. (2001). Value creation in E-business. *Strategic Management Journal*, 22(6–7), 493–520. <https://doi.org/10.1002/smj.187>
- Amshoff, B., Dülme, C., Echterfeld, J., & Gausemeier, J. (2015). Business model patterns for disruptive technologies. *International Journal of Innovation Management*, 19(3). <https://doi.org/10.1142/S1363919615400022>
- Andries, P., & Debackere, K. (2013). Business Model Innovation: Propositions on the Appropriateness of Different Learning Approaches. *Creativity and Innovation Management*, 22(4), 337–358. <https://doi.org/10.1111/caim.12033>
- Ansari, S. S., Garud, R., & Kumaraswamy, A. (2016). The disruptor's dilemma: TiVo and the U.S. television ecosystem. *Strategic Management Journal*, 37(9), 1829–1853. <https://doi.org/10.1002/smj.2442>
- Arora, S., Hofstede, F. Ter, & Mahajan, V. (2017). The implications of offering free versions for the performance of paid mobile apps. *Journal of Marketing*, 81(6), 62–78. <https://doi.org/10.1509/jm.15.0205>
- Aversa, P., Haefliger, S., & Reza, D. G. (2017). *Building a Winning Business Model Portfolio*.
- Baden-Fuller, C., & Haefliger, S. (2013). Business Models and Technological Innovation. *Long Range Planning*, 46(6), 419–426. <https://doi.org/10.1016/j.lrp.2013.08.023>

- Baden-Fuller, C., & Morgan, M. S. (2010). Business models as models. *Long Range Planning*, 43(2–3), 156–171. <https://doi.org/10.1016/j.lrp.2010.02.005>
- Bamberger, P., & Fiegenbaum, A. (1996). The role of strategic reference points in explaining the nature and consequences of human resource strategy. *Academy of Management Review*, 21(4), 926–958. <https://doi.org/10.5465/AMR.1996.9704071859>
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Barreto, I. (2010). Dynamic Capabilities: A Review of Past Research and an Agenda for the Future. *Journal of Management*, 36(1), 256–280. <https://doi.org/10.1177/0149206309350776>
- Behrens, J. T. (1997). Principles and Procedures of Exploratory Data Analysis. *Psychological Methods*, 2(2), 131–160. <https://doi.org/10.1037/1082-989X.2.2.131>
- Benner, M. J. (2010). Securities analysts and incumbent response to radical technological change: Evidence from digital photography and internet telephony. *Organization Science*, 21(1), 42–62. <https://doi.org/10.1287/orsc.1080.0395>
- Bharadwaj, A., Sawy, O. El, Pavlou, P., & Venkatraman, N. (2013). Digital business strategy: toward a next generation of insights. *MIS Quarterly*, 37(2), 471–482. <https://www.jstor.org/stable/43825919>
- Bhojraj, S., Lee, C. M. C., & Oler, D. K. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research*, 41(5), 745–774. <https://doi.org/10.1046/j.1475-679X.2003.00122.x>
- Bohnsack, R., Pinkse, J., & Kolk, A. (2014). Business models for sustainable technologies: Exploring business model evolution in the case of electric vehicles. *Research Policy*, 43(2), 284–300. <https://doi.org/10.1016/j.respol.2013.10.014>
- Bowman, D., & Gatignon, H. (1995). Determinants of Competitor Response Time to a New Product Introduction. *Journal of Marketing Research*, 32(1), 42–53. <https://doi.org/10.1177/002224379503200106>
- Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.
- Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton.

- Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications: Profound change is coming, but roles for humans remain. *Science*, 358(6370), 1530–1534. <https://doi.org/10.1126/science.aap8062>
- Cachon, G. P. (2019). A Research Framework for Business Models: What Is Common Among Fast Fashion, E-Tailing, and Ride Sharing? *Management Science*. <https://doi.org/10.1287/mnsc.2018.3275>
- Calabretta, G., Gemser, G., & Wijnberg, N. M. (2017). The Interplay between Intuition and Rationality in Strategic Decision Making: A Paradox Perspective. *Organization Studies*, 38(3–4), 365–401. <https://doi.org/10.1177/0170840616655483>
- Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance*. Quality Press.
- Casadesus-Masanell, R., & Ricart, J. E. (2010). From strategy to business models and onto tactics. *Long Range Planning*, 43(2–3), 195–215. <https://doi.org/10.1016/j.lrp.2010.01.004>
- Casadesus-Masanell, R., & Tarzijan, J. (2012). When One Business Model Isn't Enough. *Harvard Business Review*, 1–2(90), 132–137.
- Casadesus-Masanell, R., & Zhu, F. (2010). Strategies to fight ad-sponsored rivals. *Management Science*, 56(9), 1484–1499. <https://doi.org/10.1287/mnsc.1100.1199>
- Casadesus-Masanell, R., & Zhu, F. (2013). Business model innovation and competitive imitation: The case of sponsor-based business models. *Strategic Management Journal*, 34(4), 464–482. <https://doi.org/10.1002/smj.2022>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Christensen, C. M. (2006). The Ongoing Process of Building a Theory of Disruption. *Journal of Product Innovation Management*, 23(1), 39–55. <https://doi.org/10.1111/j.1540-5885.2005.00180.x>
- Christensen, C. M., McDonald, R., Altman, E. J., & Palmer, J. E. (2018). Disruptive Innovation: An Intellectual History and Directions for Future Research. *Journal of Management Studies*, 55(7), 1043–1078. <https://doi.org/10.1111/joms.12349>
- Christensen, C. M., Raynor, M., & McDonald, R. (2015). *What Is Disruptive Innovation?*

- Christensen, C. M., & Rosenbloom, R. S. (1995). Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network. *Research Policy*, 24(2), 233–257. [https://doi.org/10.1016/0048-7333\(93\)00764-K](https://doi.org/10.1016/0048-7333(93)00764-K)
- Christensen, K., Nørskov, S., Frederiksen, L., & Scholderer, J. (2017). In Search of New Product Ideas: Identifying Ideas in Online Communities by Machine Learning and Text Mining. *Creativity and Innovation Management*, 26(1), 17–30. <https://doi.org/10.1111/caim.12202>
- Davenport, T. H., & Kirby, J. (2016). Only humans need apply: winners and losers in the age of smart machines. *Harvard Business School*, August, 1–10.
- Del Giudice, M., & Straub, D. (2011). IT and entrepreneurship: An on-again, off-again love affair or a marriage? In *MIS Quarterly: Management Information Systems* (Vol. 35, Issue 4). <https://doi.org/10.2307/41409961>
- Dellermann, D., Lipusch, N., Ebel, P., & Leimeister, J. M. (2018). Design principles for a hybrid intelligence decision support system for business model validation. In: *Electronic Markets*. *Electronic Markets*, 29(3), 423–441. <https://doi.org/10.1007/s12525-018-0309-2>
- Desyllas, P., & Sako, M. (2013). Profiting from business model innovation: Evidence from Pay-As-You-Drive auto insurance. *Research Policy*, 42(1), 101–116. <https://doi.org/10.1016/j.respol.2012.05.008>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <http://arxiv.org/abs/1810.04805>
- Ebel, P., Bretschneider, U., & Leimeister, J. M. (2016). Leveraging virtual business model innovation: a framework for designing business model development tools. *Information Systems Journal*, 26(5), 519–550. <https://doi.org/10.1111/isj.12103>
- Eisenhardt, K. M., Graebner, M. E., & Sonenshein, S. (2016). Grand challenges and inductive methods: Rigor without rigor mortis. In *Academy of Management Journal* (Vol. 59, Issue 4, pp. 1113–1123). Academy of Management. <https://doi.org/10.5465/amj.2016.4004>
- Eklund, J., & Kapoor, R. (2019). Pursuing the New While Sustaining the Current: Incumbent Strategies and Firm Value During the Nascent Period of Industry Change. *Organization Science*, 30(2), 383–404. <https://doi.org/10.1287/orsc.2018.1229>
- Elnathan, D., & Kim, O. (1995). Partner selection and group formation in cooperative benchmarking. *Journal of Accounting and Economics*, 19(2–3), 345–364.

[https://doi.org/10.1016/0165-4101\(94\)00387-K](https://doi.org/10.1016/0165-4101(94)00387-K)

- Ferreira, K. J., Lee, B. H. A., & Simchi-Levi, D. (2016). Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing and Service Operations Management, 18*(1), 69–88. <https://doi.org/10.1287/msom.2015.0561>
- Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations, 7*(2), 117–140. <https://doi.org/10.1177/001872675400700202>
- Fjeldstad, Ø. D., & Snow, C. C. (2018). Business models and organization design. *Long Range Planning, 51*(1), 32–39. <https://doi.org/10.1016/j.lrp.2017.07.008>
- Foss, N. J., & Saebi, T. (2017). Fifteen Years of Research on Business Model Innovation. *Journal of Management, 43*(1), 200–227. <https://doi.org/10.1177/0149206316675927>
- Frankenberger, K., & Sauer, R. (2019). Cognitive antecedents of business models: Exploring the link between attention and business model design over time. *Long Range Planning, 52*(3), 283–304. <https://doi.org/10.1016/j.lrp.2018.05.001>
- Frankenberger, K., Weiblen, T., Csik, M., & Gassmann, O. (2013). The 4I-framework of business model innovation: A structured view on process phases and challenges. *International Journal of Product Development, 18*(3–4), 249–273. <https://doi.org/10.1504/IJPD.2013.055012>
- Gassmann, O., Frankenberger, K., & Csik, M. (2014). *The business model navigator: 55 models that will revolutionise your business*. Pearson.
- Ghezzi, A., & Cavallo, A. (2020). Agile Business Model Innovation in Digital Entrepreneurship: Lean Startup Approaches. *Journal of Business Research, 110*, 519–537. <https://doi.org/10.1016/j.jbusres.2018.06.013>
- Günzel, F., & Holm, A. B. (2013). One size does not fit all—understanding the front-end and back-end of business model innovation. *International Journal of Innovation Management, 17*(1). <https://doi.org/10.1142/S1363919613400021>
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. In *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier Inc. <https://doi.org/10.1016/C2009-0-61819-5>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning The Elements of Statistical Learning Data Mining, Inference, and Prediction, Second Edition. In *Springer series in statistics* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387->

84858-7

- Hermelo, F. D., & Vassolo, R. (2010). Institutional development and hypercompetition in emerging economies. *Strategic Management Journal*, 31(13), 1457–1473. <https://doi.org/10.1002/smj.898>
- Hill, C. W. L., & Rothaermel, F. T. (2003). The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review*, 28(2), 257–274. <https://doi.org/10.5465/AMR.2003.9416161>
- Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, 3(2), 119–131. <https://doi.org/10.1007/s40708-016-0042-6>
- Iansiti, M., & Lakhani, K. R. (2020). Competing in the Age of AI. *Harvard Business Review*. <https://hbr.org/2020/01/competing-in-the-age-of-ai>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. In *Science* (Vol. 349, Issue 6245, pp. 255–260). American Association for the Advancement of Science. <https://doi.org/10.1126/science.aaa8415>
- Kadan, O., Madureira, L., Wang, R., & Zach, T. (2012). Analysts' industry expertise. *Journal of Accounting and Economics*, 54(2–3), 95–120. <https://doi.org/10.1016/j.jacceco.2012.05.002>
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-O](https://doi.org/10.1016/0010-0285(92)90007-O)
- Kapoor, R., & Klueter, T. (2015). Decoding the adaptability-rigidity puzzle: Evidence from pharmaceutical incumbents' pursuit of gene therapy and monoclonal antibodies. *Academy of Management Journal*, 58(4), 1180–1207. <https://doi.org/10.5465/amj.2013.0430>
- Kim, S. K., & Min, S. (2015). Business Model Innovation Performance: When does Adding a New Business Model Benefit an Incumbent? *Strategic Entrepreneurship Journal*, 9(1), 34–57. <https://doi.org/10.1002/sej.1193>
- Kind, H. J., Nilssen, T., & Sørsgard, L. (2009). Business models for media firms: Does competition matter for how they raise revenue? *Marketing Science*, 28(6), 1112–1128. <https://doi.org/10.1287/mksc.1090.0514>
- Kunisch, S. (2017). Benchmarking Corporate Headquarters – Instrumental, but not Strategic?:

- Comments on Young (1998). In *Long Range Planning* (Vol. 50, Issue 1, pp. 30–32). Elsevier Ltd. <https://doi.org/10.1016/j.lrp.2016.12.007>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- Loughran, T., & Ritter, J. (2004). Why Has IPO Underpricing Changed Over Time? *Financial Management*, 33(3).
- Lucas, H. C., & Goh, J. M. (2009). Disruptive technology: How Kodak missed the digital photography revolution. *Journal of Strategic Information Systems*, 18(1), 46–55. <https://doi.org/10.1016/j.jsis.2009.01.002>
- Markides, C. (2006). Disruptive innovation: In need of better theory. In *Journal of Product Innovation Management* (Vol. 23, Issue 1, pp. 19–25). John Wiley & Sons, Ltd. <https://doi.org/10.1111/j.1540-5885.2005.00177.x>
- Martins, L. L., Rindova, V. P., & Greenbaum, B. E. (2015). Unlocking the Hidden Value of Concepts: A Cognitive Approach to Business Model Innovation. *Strategic Entrepreneurship Journal*, 9(1), 99–117. <https://doi.org/10.1002/sej.1191>
- Massa, L., Tucci, C. L., & Afuah, A. (2017). A critical assessment of business model research. In *Academy of Management Annals* (Vol. 11, Issue 1, pp. 73–104). Routledge. <https://doi.org/10.5465/annals.2014.0072>
- McDonald, R. M., & Eisenhardt, K. M. (2019). Parallel Play: Startups, Nascent Markets, and Effective Business-model Design. *Administrative Science Quarterly*, 000183921985234. <https://doi.org/10.1177/0001839219852349>
- McMullen, J. S., Shepherd, D. A., & Patzelt, H. (2009). Managerial (In)attention to Competitive Threats. *Journal of Management Studies*, 46(2), 157–181. <https://doi.org/10.1111/j.1467-6486.2008.00799.x>
- McNamara, G., Vaaler, P. M., & Devers, C. (2003). Same as it ever was: the search for evidence of increasing hypercompetition. *Strategic Management Journal*, 24(3), 261–278. <https://doi.org/10.1002/smj.295>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, January 16). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed

- Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems*. <http://arxiv.org/abs/1310.4546>
- Min, H. (2010). Artificial intelligence in supply chain management: Theory and applications. *International Journal of Logistics Research and Applications*, 13(1), 13–39. <https://doi.org/10.1080/13675560902736537>
- Murphy, S. E., Blyth, D., & Fiedler, F. E. (1992). Cognitive resource theory and the utilization of the leader's and group members' technical competence. *The Leadership Quarterly*, 3(3), 237–255. [https://doi.org/10.1016/1048-9843\(92\)90014-7](https://doi.org/10.1016/1048-9843(92)90014-7)
- Olito, F. (2020, January 16). *The rise and fall of Blockbuster*. <https://www.businessinsider.com/rise-and-fall-of-blockbuster>
- Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Wiley.
- Poole, D. L., Mackworth, A. K., & Goebel, R. (1998). *Computational intelligence : a logical approach*. Oxford University Press.
- Porter, M. E. (2001). Strategy and the Internet. *Harvard Business Review*. www.hbr.org
- Raisch, S., & Krakowski, S. (2020). Artificial Intelligence and Management: The Automation-Augmentation Paradox. *Academy of Management Review*. <https://doi.org/10.5465/2018.0072>
- Raj, M., & Seamans, R. (2019). Primer on artificial intelligence and robotics. *Journal of Organization Design*, 8(1), 11. <https://doi.org/10.1186/s41469-019-0050-0>
- Remane, G., Hanelt, A., Tesch, J. F., & Kolbe, L. M. (2017). The Business Model Pattern Database-A Tool For Systematic Business Model Innovation. *International Journal of Innovation Management*, 21(1). <https://doi.org/10.1142/S1363919617500049>
- Rietveld, J. (2018). Creating and capturing value from freemium business models: A demand-side perspective. *Strategic Entrepreneurship Journal*, 12(2), 171–193. <https://doi.org/10.1002/sej.1279>
- Riley, T. (2018, March 13). *Get ready, this year your next job interview may be with an A.I. robot*. CNBC. <https://www.cnbc.com/2018/03/13/ai-job-recruiting-tools-offered-by-hirevue-my-other-start-ups.html>
- Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence: a modern approach* (3rd

- ed.). Prentice Hall.
- Sabatier, V., Mangematin, V., & Rousselle, T. (2010). From recipe to dinner: Business model portfolios in the european biopharmaceutical industry. *Long Range Planning*, 43(2–3), 431–447. <https://doi.org/10.1016/j.lrp.2010.02.001>
- Saebi, T., Lien, L., & Foss, N. J. (2017). What Drives Business Model Adaptation? The Impact of Opportunities, Threats and Strategic Orientation. *Long Range Planning*, 50(5), 567–581. <https://doi.org/10.1016/j.lrp.2016.06.006>
- Shankar, V. (2018). How artificial intelligence (AI) is reshaping retailing. *Journal of Retailing*, 94(4), 6–11.
- Sood, A., & Tellis, G. J. (2011). Demystifying disruption: A new model for understanding and predicting disruptive technologies. *Marketing Science*, 30(2), 339–354. <https://doi.org/10.1287/mksc.1100.0617>
- Sprouts Corporate. (n.d.). *Our History - Sprouts Corporate - Natural & Organic Grocery Store*. Retrieved June 15, 2020, from <https://about.sprouts.com/about/>
- Stephan, M., Brown, D., & Erickson, R. (2017). Talent acquisition through predictive hiring. *Deloitte Insights*. <https://www2.deloitte.com/us/en/insights/focus/human-capital-trends/2017/predictive-hiring-talent-acquisition.html>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- Szopinski, D., Schoormann, T., John, T., Knackstedt, R., & Kundisch, D. (2019). Software tools for business model innovation: current state and future challenges. *Electronic Markets*, 1–26. <https://doi.org/10.1007/s12525-018-0326-1>
- Täuscher, K. (2018). Using qualitative comparative analysis and system dynamics for theory-driven business model research. *Strategic Organization*, 16(4), 470–481. <https://doi.org/10.1177/1476127017740535>
- Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, 43(2–3), 172–194. <https://doi.org/10.1016/j.lrp.2009.07.003>
- Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*. <https://doi.org/10.1016/j.lrp.2017.06.007>
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management.

- Strategic Management Journal*, 18(7), 509–533. [https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7<509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)
- Tidhar, R., & Eisenhardt, K. M. (2020). Get Rich or Die Trying... Finding Revenue Model fit Using Machine Learning and Multiple Cases. *Strategic Management Journal*, smj.3142. <https://doi.org/10.1002/smj.3142>
- Trimi, S., & Berbegal-Mirabent, J. (2012). Business model innovation in entrepreneurship. *International Entrepreneurship and Management Journal*, 8(4), 449–465. <https://doi.org/10.1007/s11365-012-0234-3>
- Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: evidence from digital imaging. *Strategic Management Journal*, 21(10-11), 1147–1161. [https://doi.org/10.1002/1097-0266\(200010/11\)21:10/11<1147::AID-SMJ128>3.0.CO;2-R](https://doi.org/10.1002/1097-0266(200010/11)21:10/11<1147::AID-SMJ128>3.0.CO;2-R)
- Tushman, M. L., & Anderson, P. (1986). Technological Discontinuities and Organizational Environments. *Administrative Science Quarterly*, 31(3), 439. <https://doi.org/10.2307/2392832>
- Wang, J. (2018, October 17). *Tesla Through the Lens of Apple - ARK Invest*. <https://ark-invest.com/analyst-research/tesla-through-the-lens-of-apple/>
- Warner, K. S. R., & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, 52(3), 326–349. <https://doi.org/10.1016/j.lrp.2018.12.001>
- Watson, G. H. (1993). *Strategic Benchmarking*. Wiley.
- Watson, G. H. (2007). *Strategic Benchmarking Reloaded with SixSigma: Improve Your Company's Performance Using Global Best Practice*. John Wiley & Sons, Inc.
- Wiggins, R. R., & Ruefli, T. W. (2005). Schumpeter's ghost: Is hypercompetition making the best of times shorter? *Strategic Management Journal*, 26(10), 887–911. <https://doi.org/10.1002/smj.492>
- Wilson, H. J., & Daugherty, P. R. (2018). Collaborative Intelligence: Humans and AI Are Joining Forces. *Harvard Business Review*. https://www.accenture.com/t00010101t000000z__w__/_acnmedia/pdf-84/accenture-collaborative-intelligence-2018.pdf
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data Mining: Practical machine*

learning tools and techniques (4th ed.). Elsevier.

- Wolfram Cox, J. R., Mann, L., & Samson, D. (1997). Benchmarking as a Mixed Metaphor: Disentangling Assumptions of Competition and Collaboration. *Journal of Management Studies*, 34(2), 285–314. <https://doi.org/10.1111/1467-6486.00052>
- Yoo, Y., Boland, R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, 23(5), 1398–1408. <https://doi.org/10.1287/orsc.1120.0771>
- Zott, C., & Amit, R. (2008). The fit between product market strategy and business model: implications for firm performance. *Strategic Management Journal*, 29(1), 1–26. <https://doi.org/10.1002/smj.642>
- Zott, C., & Amit, R. (2010). Business model design: An activity system perspective. *Long Range Planning*, 43(2–3), 216–226. <https://doi.org/10.1016/j.lrp.2009.07.004>
- Zott, C., Amit, R., & Massa, L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*, 37(4), 1019–1042. <https://doi.org/10.1177/0149206311406265>
- Zucker, L. G. (1987). Institutional theories of organization. *Annual Review of Sociology*. Vol. 13, 443–464. <https://doi.org/10.1146/annurev.so.13.080187.002303>

Appendix I: Business Model Pattern Description

The BMs of companies were replicated by assigning the 55 BMPs of Gassmann et al. (2014) to them based on the aggregated cases which consist of company descriptions. The BMPs were then categorized into meta-components following the descriptions of Günzel and Holm, (2013) and Remane et al. (2017). Table 6 gives an overview of the BMPs and their categorization.

PATTERN NAME	DESCRIPTION	META-COMPONENT
ADD-ON	The core offering is priced competitively, but there are numerous extras that drive the final price up. In the end, the customer may pay more than was originally assumed. Customers benefit from a variable offer that they can adapt to their specific needs.	Value Proposition
AFFILIATION	The focus lies in supporting others to sell products successfully, thus benefiting directly from successful transactions. Affiliates usually have some kind of pay-per-sale or pay-per-display system. The company itself gains access to a more diverse potential customer base without any additional active sales or marketing efforts.	Value Delivery
AIKIDO	Aikido allows a company to offer something diametrically opposed to the image and mindset of the competition. The novelty of the value proposition attracts the type of customer who prefer ideas or concepts that diverge from the mainstream.	Value Proposition
AUCTION	Auctioning involves selling a product or service to the highest bidder. The final price is attained at a predetermined moment or when no higher bid has been received. This allows a company to sell at the highest price acceptable to customers. The customer benefits from the opportunity to exert an influence on the price of the product.	Value Capture
BARTER	Barter is a method of exchanging goods with no transfer of money. In the business context the customer provides something of value to the sponsoring organization. The goods exchanged do not have to have a direct connection and are likely to be valued differently by each party.	Value Proposition

CASH MACHINE	According to the Cash Machine concept, the customer pays upfront for the products sold before the company has to cover the associated expenses. This results in increased liquidity that can be used to amortize debts or fund investments in	Value Creation
CROSS-SELLING	In this model, services or products from an outside business are added to the offerings, thus leveraging existing key skills and resources. In retail especially, companies can easily provide additional products and offerings that are not linked to their main focus. In this way more potential customer needs can be satisfied and additional revenue generated with relatively few changes to the existing infrastructure and assets.	Value Proposition
CROWD-FUNDING	A product, project or entire start-up is financed by a group of investors who wish to support the underlying idea, typically via the Internet. If the critical mass is achieved, the idea will be realized and investors receive special benefits, usually proportionate to the amount of money they have provided.	Value Capture
CROWD-SOURCING	The solution to a task or problem is adopted by an anonymous crowd, typically via the Internet. Contributors receive a small reward or have a chance to win a prize if their solution is chosen for production or sale. Customer interaction and inclusion can foster a positive relationship with them and subsequently increase sales and revenue for the company.	Value Creation
CUSTOMER LOYALTY	Customers are retained and loyalty assured by providing value over and above the actual product or service itself, for example through incentive-based programs. The goal is to enhance loyalty by creating an emotional connection or simply rewarding it with special offers. Customers are bound to the company voluntarily, and this protects future revenue	Value Delivery
DIGITIZATION	This pattern relies on the ability to turn existing products or services into digital versions of themselves, which thus offer advantages over tangible products, such as easier and more rapid distribution. Ideally, the digitization of a product or service should not reduce the perceived customer value.	Value Creation
DIRECT SELLING	Direct Selling refers to a scenario whereby a company's products are not sold through an intermediary but are available directly from the manufacturer or service provider. In this way, the company avoids the retail margin or any additional costs associated with the middleman. These savings can be passed on to the customer. The pattern helps to establish a uniform distribution model and the direct contact enhances customer relationships.	Value Delivery

E-COMMERCE	Traditional products or services are delivered through online channels only, thus removing costs associated with running a physical branch infrastructure. Customers benefit from greater availability and convenience, while the company is able to integrate its sales and distribution with other internal processes.	Value Delivery
EXPERIENCE SELLING	The value of a product or service is increased by an additional customer experience offered with it. This opens the door to higher customer demand and a commensurate increase in the prices charged. The customer experience needs to be adapted accordingly, for example by appropriate promotion or additional shop fittings.	Value Proposition
FLAT RATE	In this model, a single fixed fee is charged for a product or service, regardless of actual usage. The user benefits from a simple cost structure while the company benefits from a constant revenue stream.	Value Capture
FRACTIONAL OWNERSHIP	Fractional Ownership describes the sharing of a certain asset class among a group of owners. Typically, the asset is capital-intensive but is only required on an occasional basis. While the customer benefits from the owner rights, the entire capital does not have to be provided by him or her alone.	Value Creation
FRANCHISING	The franchisor owns the brand name, products and corporate identity and licenses them to independent franchisees who bear the risk of local operations. Revenue is generated as part of the franchisees' revenue and orders. The benefit for the franchisee is in the marketing of well-known brands and the availability of know-how and support.	Value Creation
FREEMIUM	The basic version of an item is offered for free in the hope of eventually persuading customers to purchase a premium version. The free offering attracts the highest volume of customers possible for the company, while revenue is generated by the (generally smaller) volume of premium customers.	Value Proposition
FROM PUSH TO PULL	This pattern describes the strategy of a company to decentralize and thus add flexibility to the company's processes in order to be more customer-focused. To respond rapidly and flexibly to new customer needs, any part of the value chain – including production or even research and development – may be affected.	Value Creation
GUARANTEED AVAILABILITY	This pattern makes the customer's needs central to decisions within the enterprise and the shaping of the value proposition. It can be applied to all aspects of the business.	Value Proposition

HIDDEN REVENUE	The logic that the income of the business depends on the users is abandoned. Instead, the main source of revenue comes from a third party, who cross finances whatever free or low-priced offering attracts the users. A very common application of this model is financing through advertisements: the customers so attracted are of value to the advertisers, who then fund the offering. This concept facilitates the concept of separation of revenue and customer.	Value Capture
INGREDIENT BRANDING	This is the inclusion of a branded ingredient originating from a different supplier into a product. The principal product is then advertised as containing the ingredient product and stressing the added value it brings to the customer. The positive association with the ingredient brand is projected on to the product and increases its attractiveness.	Value Delivery
INTEGRATOR	A company functioning on the Integrator model has command of the majority of the steps in the value adding process, including all resources and capabilities in terms of value creation. Efficiency gains, economies of scope and reduced dependency on suppliers result in a decrease in costs and may increase the stability of value creation.	Value Creation
LAYER PLAYER	A Layer Player is a specialized company limited to providing one value-adding step to different value chains. This step is typically offered within a variety of independent markets and industries. The company benefits from economies of scale and often leads to more efficient production. Furthermore, the established special expertise can result in a higher quality process	Value Creation
LEVERAGE CUSTOMER DATA	New value is created by collecting customer data and preparing it in beneficial ways for internal usage or transmission to interested third parties. Revenues are generated by either selling the data directly to others or leveraging them for the company's own purposes, e.g. to increase the effectiveness of advertising.	Value Capture
LICENSING	Here, the efforts are focused on developing intellectual property that can be licensed to other manufacturers. Thus this model relies not on the realization and utilization of knowledge in the form of creating products, but attempts to transform these intangible assets into money. Licensing gives a company the freedom to focus on research and development and allows the provision to third parties of knowledge that would otherwise be left unused.	Value Capture
LOCK-IN	Here, customers are locked into a vendor's world of products and services. Transferring custom to another vendor is impossible without incurring substantial switching costs. The Lock-in is effected either by technological mechanisms or a high level of interdependencies of products or services.	Value Creation

LONG TAIL	Rather than concentration on blockbusters, the main bulk of revenues is generated through a 'long tail' of niche products which, individually, neither demand high volumes nor allow a high margin. If a wide variety of these products is offered in sufficient amounts, the profits from the resulting accumulated small sales can add up to a significant amount.	Value Delivery
MAKE MORE OF IT	Know-how and other assets available in the company are not only used to build its own products, but are also offered to other companies. Thus slack resources are used to create additional revenue besides those generated directly by the company's core value proposition.	Value Capture
MASS CUSTOMISATION	Customizing products through mass production once seemed to be an impossible endeavor, but this has now changed with the development of modular products and production systems that enable efficient individualization of products. As a result, individual customer needs can be met under mass production conditions and at competitive prices.	Value Creation
NO FRILLS	No Frills value creation focuses on the necessary minimum to deliver the core value proposition of a product or service, which will thus typically be very basic. Cost savings are shared with the customer, usually resulting in a customer base with lower purchasing power or purchasing willingness.	Value Proposition
OPEN BUSINESS	In Open Business models, collaboration with partners in the ecosystem becomes a central source of value creation. Companies pursuing an Open Business model actively search for novel ways of working together with suppliers, customers or complementors to open up and extend their business.	Value Creation
OPEN SOURCE	In Open Source software engineering, the source code of a software product is not proprietary, but is made freely accessible for anyone. Generally, this could be applied to any technological details of any product. Others can contribute to the product, but also use it freely as solely user. Money is typically earned with services that are complimentary to the product, such as consulting and support.	Value Creation
ORCHESTRATOR	In this model, the company's focus is on the core competencies within the value chain. The other segments of the value chain are outsourced and actively coordinated. This allows the company to reduce costs and to benefit from suppliers' economies of scale. The focus on core competencies can enhance performance.	Value Creation

PAY PER USE	In this model, the actual usage of a service or product is metered, that is to say, the customer pays on the basis of what is effectively consumed. In this way the company attracts customers who wish to benefit from the additional flexibility, which might be priced higher.	Value Capture
PAY WHAT YOU WANT	The buyer pays any desired amount for a given commodity, sometimes even zero. In some cases, a minimum floor price may be set, and/or a suggested price may be indicated as guidance for the buyer. The attraction for the customer is the ability to influence the price, while the seller benefits from a larger number of customers.	Value Proposition
PEER TO PEER	This model (often abbreviated as P2P) is based on a cooperation among individuals belonging to an homogeneous group. The organising company offers a meeting point, normally an online database and communication service, which connects these individuals. Examples of transactions are the offering of personal items for rent, provision of certain products or services, or the sharing of information and experiences.	Value Creation
PERFORMANCE-BASED CONTRACTING	The price of a product here is based not on its physical value, but on the performance or valuable outcome it delivers in the form of a service. Performance-based contractors are often strongly integrated into the value creation process of their customers. Special expertise and economies of scale result in lower production and maintenance costs, that can be passed on to the customer.	Value Capture
RAZOR AND BLADE	The basic product is cheap or given away for free, while the consumables are expensive and sold at high margins. The price of the initial product lowers customers' barriers to purchase, while the subsequent recurring sales cross-finance it. Usually, these products are technologically bound to each other to anchor this effect more firmly.	Value Capture
RENT INSTEAD OF BUY	Here, instead of buying a product, the customer rents it. This reduces the capital typically needed to gain access to the product. The company itself benefits from higher profits on each product, as it is paid for the duration of the rental period. Both parties benefit from greater efficiency in product utilisation, given that time of non-usage, which unnecessarily ties capital down, is reduced.	Value Capture
REVENUE SHARING	Revenue Sharing refers to the practice of sharing revenues with one's stakeholders, such as complementors or even rivals. One party obtains a share of the revenue from another that benefits from the increased value of its customer base.	Value Capture

REVERSE ENGIN-NEERING	This pattern refers to obtaining a competitor's product, taking it apart and using the information obtained to produce a similar or compatible product. Because no great investment in research or development is necessary, these products can be offered at a lower price than the original one.	Value Creation
REVERSE INNOVATION	Simple inexpensive products that have been developed within and for emerging markets are also sold in industrial countries. The adjective 'reverse' here refers to the difference from the usual process whereby new products are developed in industrial countries and adapted to fit emerging market needs.	Value Creation
ROBIN HOOD	The same product or service is made available to 'the rich' at a much higher price than to 'the poor', so that the bulk of the profits are generated from the wealthy customer base. While serving 'the poor' is not profitable per se, it creates economies of scale that other providers cannot achieve. Additionally, it has a positive effect on the company's image.	Value Delivery
SELF-SERVICE	Part of the value creation of the service or product is transferred to the customer in exchange for a lower price. This is particularly suited for process steps that add relatively little perceived value for the customer, but in fact incur high costs. Customers benefit from efficiency and time savings. Efficiency may even be increased, as in some cases the customer is able to execute a value adding step more quickly and in a more target oriented manner than the company.	Value Creation
SHOP IN SHOP	Instead of opening new branches, the company finds a partner whose branches can profit from integrating its offerings, resulting in effect in a small shop within another shop (a win-win situation). The hosting store can benefit from a larger number of customers and a constant revenue in the form of rent, while the hosted company gains access to cheaper resources such as space, location or workforce.	Value Creation
SOLUTION PROVIDER	Instead of opening new branches, the company finds a partner whose branches can profit from integrating its offerings, resulting in effect in a small shop within another shop (a win-win situation). The hosting store can benefit from a larger number of customers and a constant revenue in the form of rent, while the hosted company gains access to cheaper resources such as space, location or workforce.	Value Proposition
SUBSCRIPTION	The customer pays a regular fee, typically on a monthly or annual basis, to gain access to a product or service. While customers mostly benefit from lower usage costs and general service availability, the company generates a more steady income stream.	Value Capture

SUPERMARKET	large variety of readily available products and accessories under one roof. Generally, the assortment of products is large but the prices are kept low. More customers are attracted to the wide range of goods on offer, while economies of scope yield advantages for the company.	Value Proposition
TARGET THE POOR	Here, the product or service offered targets the customer positioned at the base of the pyramid rather than the premium customer. The customers with lower purchasing power benefit from affordable products. While the company generates small profits with each product sold, it benefits from the higher sales numbers usually associated with the scale of the customer base.	Value Delivery
TRASH TO CASH	Used products are collected and either sold in other parts of the world or transformed into new products. The profit scheme is essentially based on low-to-no purchase prices. Resource costs for the company are practically eliminated, while the supplier's waste disposal is either provided free of charge or with reduced associated costs. This pattern also addresses customers' potential environmental awareness ideals.	Value Creation
TWO-SIDED MARKET	A Two-sided Market facilitates interactions between multiple interdependent groups of customers. The value of the platform increases as more groups or individual members of each group use it. The two sides frequently come from disparate groups, for example businesses on the one hand and private interest groups on the other.	Value Proposition
ULTIMATE LUXURY	This pattern describes the strategy of a company that concentrates on the upper end of society's pyramid, whereby it can distinguish its products or services strongly from others. High standards of quality or exclusive privileges are the main focus to attract this kind of customers. The investments necessary to achieve differentiations are met by the relatively high prices that can be charged and which generally allow very high margins.	Value Delivery
USER DESIGN	In this pattern, the customer is both the manufacturer and the consumer. As an example, an online platform provides the customer with the necessary support to design and merchandise the product, e.g. product design software, manufacturing services, or an online shop to sell the product. Thus, the company function is limited to supporting its customers in their undertakings and so benefits from their creativity. The customer benefits from the opportunity to realise entrepreneurial ideas without having to establish the necessary infrastructure. Revenue is then generated by the actual sales.	Value Proposition
WHITE LABEL	A White Label producer allows other companies to distribute its goods under their brand name, which thus appear to be made by them. The same product or service is often sold by multiple marketers under different brands. In this way various customer segments can be satisfied with the same product.	Value Creation

Table 6: BMP Description Overview

Reprinted from Gassmann et al. (2014)

Appendix II: F₁-Scores Original ML Model

To evaluate the accuracy of the original ML model, the corresponding F1-scores were analyzed, implicating an overall performance of worse than chance.

Pattern Name	It. 1	It. 2	It. 3	It. 4	It. 5	It. 6	It. 7	It. 8	It. 9
Auction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Robin Hood	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Add-On	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Crowdsourcing	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Performance-based Contracting	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Supermarket	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50
Cash Machine	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00
Revenue Sharing	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Reverse Innovation	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00
Mass Customisation	0.00	0.00	0.00	0.50	0.67	0.00	0.00	0.00	0.00
Freemium	0.00	0.00	0.00	0.00	0.00	1.00	0.67	0.00	0.00
Customer Loyalty	0.33	0.00	0.50	0.80	0.00	0.57	0.00	0.67	0.00
Ingredient Branding	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00
User Design	0.00	0.00	0.00	0.67	0.00	0.00	1.00	0.00	0.67
Reverse Engineering	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00
Cross-Selling	0.00	0.67	0.67	0.00	0.67	0.40	0.00	0.00	0.80
Orchestrator	0.00	0.00	0.40	0.00	0.67	0.00	0.00	0.50	0.00
From Push to Pull	0.67	0.67	0.40	0.00	0.50	0.00	0.00	0.00	0.00
Make More of It	1.00	0.67	0.67	0.00	1.00	0.00	0.00	0.00	0.00
E-Commerce	0.67	0.50	0.57	0.55	0.50	0.00	0.00	0.57	0.86
Rent Instead of Buy	0.67	0.80	0.00	0.00	0.40	0.67	0.67	0.00	0.67
Open Source	0.80	0.00	0.00	0.80	0.00	0.00	1.00	1.00	1.00
Franchising	0.00	1.00	0.67	1.00	0.00	0.67	1.00	0.00	0.00
Solution Provider	1.00	0.75	0.86	0.00	0.67	1.00	0.00	0.67	0.00
Ultimate Luxury	0.86	0.67	1.00	0.50	0.00	0.67	0.00	0.50	0.50
Open Business	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	1.00
Leverage Customer Data	0.00	0.67	0.67	0.00	0.00	1.00	0.00	0.67	1.00
Subscription	0.50	0.86	0.00	0.00	0.67	0.80	0.80	0.00	0.67
Integrator	0.75	0.00	0.50	0.00	0.80	0.67	0.75	0.89	0.67
Guaranteed Availability	0.40	0.80	0.00	0.00	0.67	0.80	0.80	0.00	1.00
Hidden Revenue	0.00	0.40	0.00	1.00	1.00	1.00	0.00	1.00	0.80
Flat Rate	0.67	0.86	0.00	0.00	1.00	0.80	0.80	0.00	1.00
Self-Service	0.67	0.67	1.00	0.67	0.57	0.80	0.67	0.40	0.50
Shop In Shop	0.50	0.67	0.67	0.75	1.00	0.67	0.40	0.00	0.86
Target the Poor	0.33	0.67	1.00	0.67	0.67	0.86	0.75	0.50	0.75
Layer Player	1.00	0.50	0.57	0.00	0.80	0.67	1.00	0.67	0.00
No Frills	0.80	1.00	1.00	0.67	0.50	0.50	0.33	0.67	0.86

Experience Selling	0.50	0.40	0.89	0.77	0.91	0.94	0.50	0.73	0.67
Long-Tail	0.80	0.75	0.50	0.67	1.00	0.33	1.00	0.00	0.80
Two-sided Market	0.80	0.80	0.67	0.40	0.89	0.80	0.00	0.80	0.80
Licensing	0.67	0.00	1.00	0.67	0.80	0.80	0.00	0.80	0.86
Direct Selling	0.67	0.67	0.92	0.67	0.67	0.70	0.83	0.92	1.00
Peer to Peer	1.00	0.89	1.00	0.50	1.00	1.00	0.00	1.00	0.80
Digitization	1.00	1.00	0.80	1.00	0.91	1.00	0.91	0.89	0.92
Pattern Name	It. 10	It. 11	It. 12	It. 13	It. 14	It. 15	Mean		
Auction	0.00	0.00		0.00	0.00	0.00	0.00	0.07	
Robin Hood	0.00	0.00		0.00	0.00	0.00	0.00	0.07	
Add-On	0.67	0.00		0.00	0.00	0.00	0.00	0.11	
Crowdsourcing	0.00	0.00		0.00	0.67	1.00	0.00	0.11	
Performance-based Contracting	0.00	0.00		0.00	0.00	1.00	0.00	0.13	
Supermarket	0.00	0.67		0.50	0.00	0.00	0.50	0.14	
Cash Machine	0.00	0.00		1.00	0.00	0.00	0.00	0.18	
Revenue Sharing	0.00	0.00		0.67	0.00	1.00	0.00	0.18	
Reverse Innovation	0.00	0.67		0.00	1.00	0.00	0.67	0.18	
Mass Customisation	0.00	0.00		0.00	0.00	1.00	1.00	0.21	
Freemium	0.00	0.00		0.50	0.00	0.50	1.00	0.24	
Customer Loyalty	0.00	0.67		0.50	0.00	0.00	0.40	0.30	
Ingredient Branding	0.00	1.00		0.00	1.00	0.00	0.00	0.33	
User Design	0.00	1.00		0.00	0.00	1.00	1.00	0.36	
Reverse Engineering	0.00	0.00		0.67	0.00	0.67	0.00	0.36	
Cross-Selling	0.67	0.86		0.00	0.00	0.00	0.67	0.36	
Orchestrator	0.67	0.67		0.00	1.00	1.00	0.67	0.37	
From Push to Pull	0.50	0.00		0.67	0.67	1.00	1.00	0.40	
Make More of It	0.80	0.00		0.00	1.00	0.67	0.67	0.43	
E-Commerce	0.00	0.67		1.00	0.50	0.67	0.00	0.47	
Rent Instead of Buy	1.00	0.50		0.50	0.67	0.67	0.00	0.48	
Open Source	0.00	1.00		0.29	0.67	0.67	0.00	0.48	
Franchising	1.00	1.00		0.00	0.00	0.50	0.67	0.50	
Solution Provider	0.57	0.00		0.29	0.67	0.57	0.67	0.51	
Ultimate Luxury	0.67	1.00		0.67	0.67	0.00	0.67	0.56	
Open Business	0.00	1.00		1.00	0.00	1.00	0.50	0.57	
Leverage Customer Data	0.00	1.00		0.67	1.00	0.89	1.00	0.57	
Subscription	0.67	0.67		0.75	1.00	0.57	0.67	0.57	
Integrator	0.80	1.00		0.80	0.73	0.00	0.67	0.60	
Guaranteed Availability	0.50	1.00		0.67	1.00	0.86	0.57	0.60	
Hidden Revenue	0.67	0.50		0.80	0.67	0.80	0.67	0.62	
Flat Rate	0.67	1.00		0.50	1.00	0.67	0.67	0.64	
Self-Service	0.67	0.50		0.33	0.29	1.00	1.00	0.65	
Shop In Shop	0.80	0.86		1.00	0.00	1.00	0.57	0.65	
Target the Poor	0.50	0.86		0.33	0.40	0.75	0.75	0.65	
Layer Player	0.50	0.91		0.75	0.86	0.57	1.00	0.65	
No Frills	0.67	1.00		0.40	0.40	0.67	0.75	0.68	
Experience Selling	0.67	0.89		0.67	0.75	0.50	0.77	0.70	

Long-Tail	0.67	1.00	0.80	1.00	0.89	0.40	0.71
Two-sided Market	0.89	0.75	0.86	1.00	0.83	0.57	0.72
Licensing	0.80	1.00	1.00	1.00	1.00	0.80	0.75
Direct Selling	0.67	0.86	0.89	0.83	1.00	0.67	0.80
Peer to Peer	1.00	0.67	0.80	0.67	0.91	0.80	0.80
Digitization	0.73	0.93	1.00	0.89	0.96	0.67	0.91

Table 7: F1-Score Results of K-Fold Cross Validation, Original Model

Appendix III: Detailed Elaboration on EDA

The EDA was conducted to gather first insights whether BMPs across firms differ which could indicate learning opportunities for BMI. Thus, the 256 firms of the OD were re-evaluated and segmented into industries. The distribution of the five most detected BMPs across industries, according to the proportion of firms using the BMP within each industry, is depicted in Figure 8. When dissecting the allocation of BMPs across industries, it becomes evident that the most utilized pattern '*Solution Provider*' is used in each of them. The highest employment is within the industry sectors Health Care, Energy, and Industrials with 91%, 89% and 81% of companies within the industries employing it, respectively. The lowest utilization of the BMP is within the Consumer Staples (14% of companies) and Consumer Discretionary (15% of companies) industry sectors.

Digitization is utilized by 100% of the companies within the industry sector Communication Services and by 80% of the companies within Information Technology. However, it was not allocated to a firm within the Energy, Materials, and Consumer Staples sectors. Being employed by 76 companies in total, *Licensing* is also less equally distributed across industry sectors. Despite its allocation to 75% of firms within the Communication Services as well as 52% within the Information Technology and 35% in the Health Care sector, it was not employed by companies of the Utilities and Real Estate sectors and only by 3% of firms within the Financials sector.

Increasing granularity by investigating the 24 industry groups according to the GICS adds additional insights. For example, the BMP *Cash Machine* describes the value creation process of charging the payment to a customer before delivering the value. In the meantime, the companies can profit from the increased liquidity. 100% of the industry groups Banks and Diversified Financials use this BMP and 86% of insurance firms. Thus, 94% of the firms within the Financials industry sector employ the *Cash Machine* BMP resulting in only two that do not.

Appendix IV: Results of Upstart Allocation

The allocation of BMPs were used for the Multi-Case Theory Building and to improve train and test the dataset as depicted in the following figure.

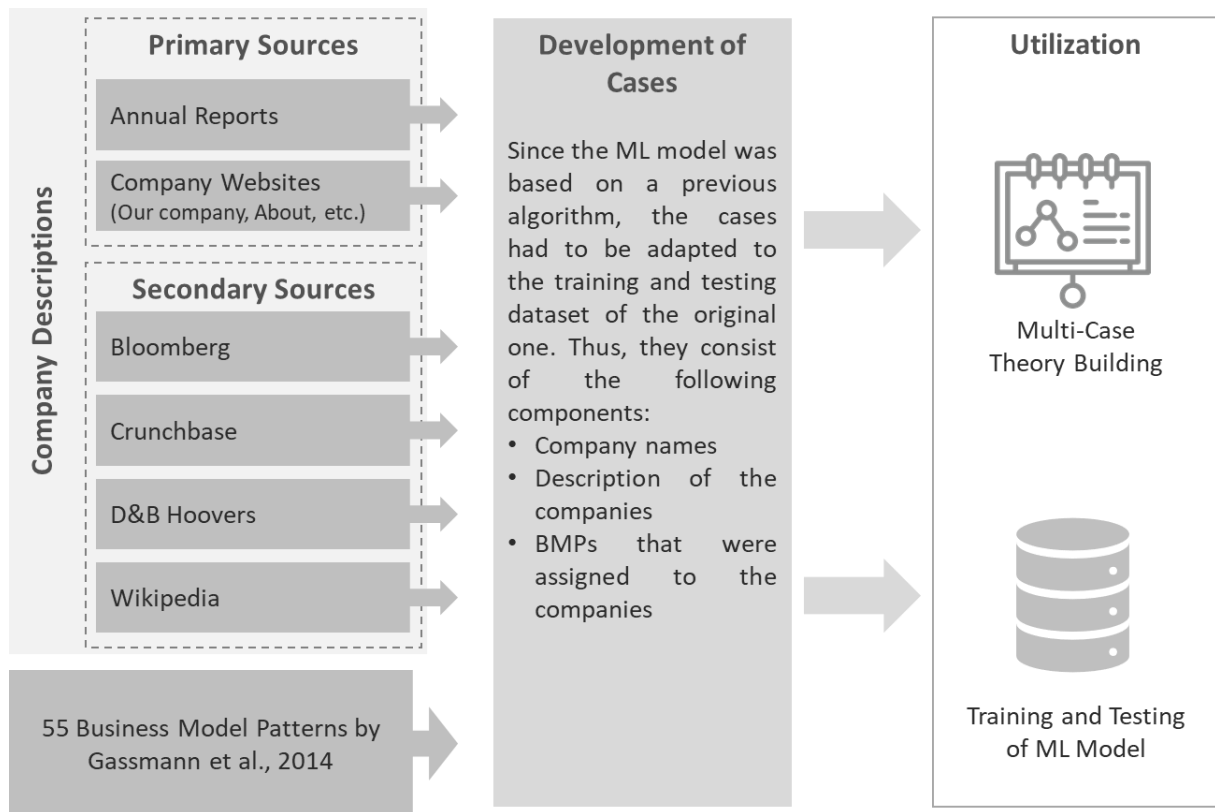


Figure 14: Process of Compiling the Cases and Utilization

Beneath table includes the allocated BMPs of above described process.

COMPANY	CATEGORY 1	CATEGORY 2
10X GENOMICS	Solution Provider, Leverage Customer Data, Experience Selling, Digitization, Lock-In, Direct Selling, Affiliation, Fractional Ownership	Razor & Blade, Aikido
AMYRIS BIOTECHNOLOGIES	Fractional Ownership, Revenue Sharing, Affiliation	From push to pull, Aikido
ATHENE HOLDING LTD	Cash Machine, Subscription, Revenue Sharing, Direct Selling, Affiliation, Aikido, Experience Selling, Orchestrator, Licensing, Direct Selling	Reverse Engineering, Ingredient Branding, Affiliation
BEYOND MEAT	Aikido, Experience Selling, Orchestrator, Licensing, Direct Selling	Reverse Engineering, Ingredient Branding, Affiliation

BLOOM ENERGY	Subscription, Layer Player, Guaranteed Availability, Direct Selling, Affiliation	Target the poor, Lock-In
BLUE APRON	Subscription, Self-Service, Cross-Selling, Experience Selling, Add-On, Orchestrator, Leverage Customer Data, Direct Selling, E-commerce, Digitization, Customer Loyalty	
BRIGHAM MINERALS	Performance-based contracting, Rent instead of buy	Leverage Customer Data
CARVANA CO	Two-Sided Market, Experience Selling, Solution Provider, Digitization, Self-Service, From trash to cash, Leverage Customer Data, E-Commerce, From push to pull, Add-On	Integrator, Customer Loyalty, Direct Selling, Supermarket
CHICKEN SOUP FOR THE SOUL ENTERTAINMENT	Digitization, Subscription, Hidden Revenue, Licensing, Experience Selling	Flat Rate, Guaranteed Availability, Open Business, Aikido
CONIFER HOLDINGS	Cash Machine, Subscription, Affiliation	
DROPBOX	Digitization, Guaranteed Availability, Peer to peer, Freemium, Subscription, Ingredient Branding, Experience Selling, Leverage Customer Data	Open Business, Self-Service
ELEVATE CREDIT	Leverage Customer Data, Direct Selling, Cash Machine, Target the poor, Digitization, Licensing	Experience Selling, Customer Loyalty, Barter
ETSY	Two-Sided Market, Revenue Sharing, Add-On, Experience Selling, E-Commerce, Leverage Customer Data, Make more of it	Self-Service, Supermarket
EVENTBRITE	Two-Sided Market, E-Commerce, Solution Provider, Experience Selling, Hidden Revenue, Lock-In, Self-Service, Affiliation, Leverage Customer Data, Subscription	Integrator
EVERQUOTE	Two-Sided Market, Experience Selling, Leverage Customer Data, Revenue Sharing, E-Commerce, Make more of it, Self-Service, Auction, Digitization, Hidden Revenue	Customer Loyalty
EVOLENT HEALTH INC	Solution Provider, Two-Sided Market, Leverage Customer Data, Digitization, Integrator, Direct Selling	Experience Selling, Aikido
FITBIT	Leverage Customer Data, Digitization, Experience Selling, Direct Selling, Two-Sided Market, Cross-Selling, Customer Loyalty, E-commerce, Freemium, Orchestrator, Affiliation	Make more of it
HUBSPOT	Digitization, Make more of it, Two-Sided Market, Self-Service, Freemium, Subscription, Revenue Sharing, Experience Selling, Affiliation, Add-On, Cross-Selling	Leverage Customer Data, Lock-In, Solution Provider
HYRECAR	Two-Sided Market, Peer to peer, Self-Service, Experience Selling, Revenue Sharing	
LENDING CLUB	Peer to peer, Two-Sided Market, Revenue Sharing, E-Commerce, Digitization, Leverage Customer Data	Make more of it, Experience Selling, Crowdfunding
LIBERTY OILFIELD SERVICES INC	Solution Provider, Performance-based contracting, Direct Selling, Orchestrator	Leverage Customer Data

LYFT	Peer to Peer, Digitization, Two-Sided Market, Leverage Customer Data, Target the poor, Experience Selling, Revenue Sharing, Open Business, Guaranteed Availability, Customer Loyalty, Subscription, Make more of it	Affiliation, Solution Provider
MATCH GROUP	Digitization, Freemium, Subscription, Hidden Revenue, Leverage Customer Data, Pay Per Use, Peer to peer	Experience Selling
MONGODB	Digitization, Subscription, Open Source, Layer Player, Make more of it, Freemium, Direct Selling, Leverage Customer Data, Performance-based contracting, Add-On, Solution Provider	Guaranteed Availability
NEW HOME CO	Leverage Customer Data, Experience Selling, Make more of it, Revenue Sharing, Direct Selling, Affiliation	Cash Machine
OPORTUN	Leverage Customer Data, Target the poor, Experience Selling, Guaranteed Availability, Digitization, Customer Loyalty	Lock-In
QUANTERIX CORP	Solution Provider, Reverse Engineering, Direct Selling, Affiliation, Open Business, Orchestrator, Make more of it	Razor & Blade
REDFIN CORP	Two-Sided Market, Digitization, Affiliation, Leverage Customer Data	Solution Provider, Experience Selling
ROKU INK	Digitization, Two-Sided Market, Experience Selling, Leverage Customer Data, Freemium, Hidden Revenue, Solution Provider, Ingredient Branding, Open Source, Licensing, Customer Loyalty, Revenue Sharing, Target the poor, Affiliation, Orchestrator, No Frills	Revenue sharing, Aikido, Lock-In, Make more of it
SHAKE SHACK	Experience Selling, Customer Loyalty, Digitization, Self-Service, Direct Selling, Open Business, Shop in Shop, Licensing, Cross-Selling	Aikido
SHIFTPIXY INC	Two-Sided Market, Digitization, Leverage Customer Data, Performance-based contracting, Solution Provider	Revenue Sharing, Self-Service
SHUTTERSTOCK INC	Two-Sided Market, Peer to Peer, Digitization, Licensing, Leverage Customer Data, Ultimate Luxury, Cross-Selling, E-Commerce, Experience Selling, Direct Selling	Revenue Sharing, Make more of it
SMART SAND	Integrator, Solution Provider, No Frills, Rent instead of buy	Performance-based contracting, Direct Selling
SMILEDIRECTCLUB	Aikido, Direct Selling, Integrator, Experience Selling, Subscription, Cross-Selling, Leverage Customer Data, Experience Selling	Mass customization, Open Business, Self-Service
SNAP INC.	Experience Selling, Digitization, Two-Sided Market, Hidden Revenue, Affiliation, Leverage Customer Data, Make more of it, Open Business, Solution Provider, Licensing	From push to pull, Affiliation
SOLID BIOSCIENCES INC	Open Business, Solution Provider, Licensing	From push to pull, Affiliation

SONOS	Digitization, Experience Selling, Open Business, Affiliation, E-Commerce, Orchestrator	
SPROUT SOCIAL	Digitization, Leverage Customer Data, Experience Selling, Layer Player, Customer Loyalty, Self-Service, Subscription, Add-On	Lock-In, Freemium
SPROUTS FARMERS MARKET	Supermarket, Experience Selling, Customer Loyalty, No Frills, Orchestrator	Robin Hood, Revenue Sharing, Aikido
SQUARE	Solution Provider, Digitization, Make more of it, Self-Service, Peer to Peer, Freemium, Subscription	Lock-In
SUNNOVA ENERGY	Affiliation, Rent instead of buy, Subscription, Make more of it	Open Business, Experience Selling
SUNRUN INC	Subscription, Revenue Sharing, Direct Selling, Target the poor, Experience Selling, Customer Loyalty, Guaranteed Availability, Affiliation	Rent instead of buy, Solution Provider, Orchestrator, Lock-In
TESLA	Direct Selling, Ultimate Luxury, Experience Selling, Cross-Selling, Integrator, Fractional Ownership, Leverage Customer Data	Digitization, Layer Player, Aikido, From trash to cash, Open Business
THE TRADE DESK	Self-Service, Digitization, Leverage Customer Data, Layer Player, Customer Loyalty, Solution Provider, Open Source, Experience Selling, Two-Sided Market, Lock-In, Cross-Selling	Auction
TWITTER	Digitization, Open Source, Experience Selling, Two-Sided Market, Leverage Customer Data, Hidden Revenue	Affiliation, Auction, Make more of it, Licensing
UPWORK	Two-Sided Market, Revenue Sharing, Leverage Customer Data, Digitization, Make more of it, Direct Selling	Self-Service
VERITONE	Open Source, Licensing, Two-Sided Market, Digitization, Leverage Customer Data, Make more of it, Open Business, Self-Service, Performance-based contracting, Direct Selling, Affiliation, Customer Loyalty	Integrator
WORKDAY	Solution Provider, Digitization, Experience Selling, Leverage Customer Data, Layer Player, Subscription, Direct Selling, Open Business	Open Source
ZOOM VIDEO COMMUNICATIONS	Digitization, Experience Selling, Two-Sided Market, Freemium, Subscription, Direct Selling, E-Commerce	Peer to Peer, Open Business
ZYNGA	Digitization, Hidden Revenue, Affiliation, Cross-Selling, Two-Sided Market, Revenue Sharing, Customer Loyalty, Add-On, Leverage Customer Data, Experience Selling, Lock-In	Peer to Peer, Barter

Table 8: Resulting Allocation of BMPs to Upstarts

Appendix V: F1-Scores Final ML Model

Beneath Table 9 depicts the results of each testing iteration of the k-fold-cross-validation for the final ML model. It excludes the BMPs which were not identifiable (F_1 -score = 0). As argued in chapter 3.1., the resulting model also excluded the six BMPs with a F_1 -score < 0.18.

Pattern Name	It. 1	It. 2	It. 3	It. 4	It. 5	It. 6	It. 7	It. 8	It. 9	It.10	It.11	Mean
Auction	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.29	0.00	0.00	0.00	0.10
Mass Customization	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.10
Crowdfunding	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
Reverse Engineering	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	1.00	0.00	0.17
Razor and Blade	0.00	0.00	1.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.17
From Push to Pull	0.67	0.67	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
Integrator	0.80	0.80	0.67	0.33	0.00	0.40	0.44	0.67	0.00	0.00	0.00	0.33
Ingredient Branding	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.67	0.67	0.00	0.33
Supermarket	0.00	0.00	0.67	0.67	0.00	0.67	0.67	0.00	0.00	0.00	0.80	0.35
Trash to Cash	1.00	1.00	0.00	0.00	1.00	1.00	0.67	0.00	0.00	0.00	0.00	0.37
Fractional Ownership	1.00	1.00	1.00	0.00	0.00	0.67	1.00	0.00	0.00	0.00	0.00	0.37
Add-On	0.57	0.57	0.67	0.33	0.40	0.22	0.57	0.67	0.00	0.00	0.29	0.37
Layer Player	0.75	0.75	0.40	0.00	0.40	0.00	0.67	0.50	0.67	0.80	0.00	0.42
Target the Poor	0.00	0.00	0.50	0.80	0.00	1.00	0.00	0.67	0.50	0.67	0.50	0.46
Guaranteed Availability	0.50	0.50	0.33	0.50	0.40	0.00	0.50	1.00	0.33	0.55	0.57	0.47
E-Commerce	0.40	0.40	0.29	0.50	0.25	0.67	0.62	0.55	0.67	0.57	0.67	0.52
Peer to Peer	0.50	0.50	0.44	0.73	0.62	0.33	0.60	0.60	0.71	0.67	0.33	0.55
Performance-based Contracting	0.60	0.60	0.33	0.00	0.67	0.33	0.80	1.00	1.00	0.86	0.00	0.56
Orchestrator	0.33	0.33	0.80	0.33	0.67	0.75	0.50	0.67	0.67	0.44	0.50	0.57
Rent Instead of Buy	0.50	0.50	0.00	1.00	0.33	0.67	0.86	0.67	1.00	0.67	0.00	0.57
Customer Loyalty	0.62	0.62	0.63	0.78	0.31	0.50	0.33	0.75	0.55	0.77	0.63	0.58
Make More of It	0.86	0.86	0.64	0.44	0.64	0.67	0.50	0.76	0.61	0.36	0.48	0.59
Shop In Shop	1.00	1.00	1.00	0.00	0.00	0.67	0.00	1.00	1.00	0.67	0.67	0.60
Freemium	0.67	0.67	0.60	0.50	0.57	0.50	0.80	0.60	0.83	0.40	0.55	0.60
Cross-Selling	0.50	0.50	0.73	0.44	0.80	0.29	0.50	0.77	0.80	0.80	0.67	0.63
Hidden Revenue	0.80	0.80	0.80	0.80	0.00	0.86	0.86	0.67	0.50	0.80	0.22	0.63
Lock-In	0.73	0.73	0.80	0.63	0.60	0.44	0.62	0.77	0.75	0.73	0.44	0.65
Cash Machine	0.67	0.67	0.00	0.67	0.67	1.00	1.00	0.50	0.33	1.00	0.80	0.66
Open Source	0.89	0.89	0.80	0.80	0.29	0.67	1.00	0.89	0.67	0.67	0.00	0.67
Revenue Sharing	0.84	0.84	0.80	0.76	0.67	0.55	0.70	0.77	0.53	0.36	0.73	0.67
Subscription	0.55	0.55	0.55	0.50	0.96	0.73	0.73	0.63	0.72	0.81	0.60	0.68
Solution Provider	0.77	0.77	0.70	0.80	0.53	0.86	0.72	0.59	0.63	0.67	0.67	0.69
Aikido	0.75	0.75	0.67	0.73	0.33	0.91	0.67	1.00	0.33	0.92	0.67	0.70

Direct Selling	0.64	0.64	0.64	0.87	0.67	0.69	0.67	0.74	0.77	0.69	0.62	0.70
Licensing	0.57	0.57	1.00	0.86	0.67	0.80	0.00	0.75	0.80	0.77	0.86	0.71
Affiliation	0.71	0.71	0.82	0.74	0.58	0.70	0.80	0.76	0.71	0.78	0.67	0.73
Open Business	0.80	0.80	0.75	0.67	0.43	0.91	0.67	0.80	0.57	0.87	0.86	0.73
Self-Service	0.82	0.82	0.93	0.86	0.86	0.85	0.74	0.80	0.70	0.75	0.79	0.81
Two-sided Market	0.79	0.79	0.79	0.87	0.67	0.93	0.86	0.95	0.75	0.86	0.67	0.81
Leverage Customer Data	0.90	0.90	0.87	0.83	0.78	0.93	0.84	0.91	0.84	0.76	0.88	0.85
Experience Selling	0.82	0.82	0.88	0.82	0.90	0.80	0.88	0.96	0.93	0.88	0.96	0.88
Digitization	0.90	0.90	0.88	0.91	0.89	0.84	0.86	0.90	0.85	0.94	0.88	0.89
Mean	0.58	0.58	0.58	0.51	0.42	0.57	0.58	0.61	0.51	0.57	0.43	0.54

Table 9: F1-Score Results of K-Fold Cross Validation, Improved Model

Appendix VI: Algorithm Training and Testing Data

In order to draw up the dataset for training and testing the ML model and to base the MCTB on, company descriptions were compiled as cases according to the illustration in Figure 14. Due to the sheer size of these cases, an inclusion in this document was not possible. Thus, the company descriptions for the 50 upstarts can be accessed under the following link:

https://www.dropbox.com/s/z3p97i46dxwqlzz/Luca%20Lenz_Dissertation_Upstart%20Cases.docx?dl=0

Further, for the S&P 500 dataset, the annual reports were evaluated and compiled in a similar fashion. Including the annual reports in one document was not feasible and thus, these company descriptions are split in three versions, accessible via following links:

1. https://www.dropbox.com/scl/fi/pswudk33mh97ju4c0gkw6/S-P500_Company-Data-Extraction_v1.docx?dl=0&rlkey=dzojudbjvhgisfjyc8g2ixfjq
2. https://www.dropbox.com/scl/fi/zkwebuybvqx3y3kbwkzu/S-P500_Company-Data-Extraction_v2.docx?dl=0&rlkey=rpncanvkqp87n056921q205k9
3. https://www.dropbox.com/scl/fi/6eiwh85thaquby4yk4qzq/S-P500_Company-Data-Extraction_v3.docx?dl=0&rlkey=kkz648qhx1m7271dcqes6v0jf

The subsequent dataset that resulted and which is comprised of the predicted BMPs, additional data on the constituents as well as the upstarts, can be accessed via following link:

https://www.dropbox.com/scl/fi/e8m9vkztqwwcg92m5jucv/Luca-Lenz_Dissertation_Final-Dataset.xlsx?dl=0&rlkey=yjov0ytdm7wh3c57u5ly438eb

Appendix VII: Customer Journey

Following figure describes a possible customer journey for users of the benchmarking tool.

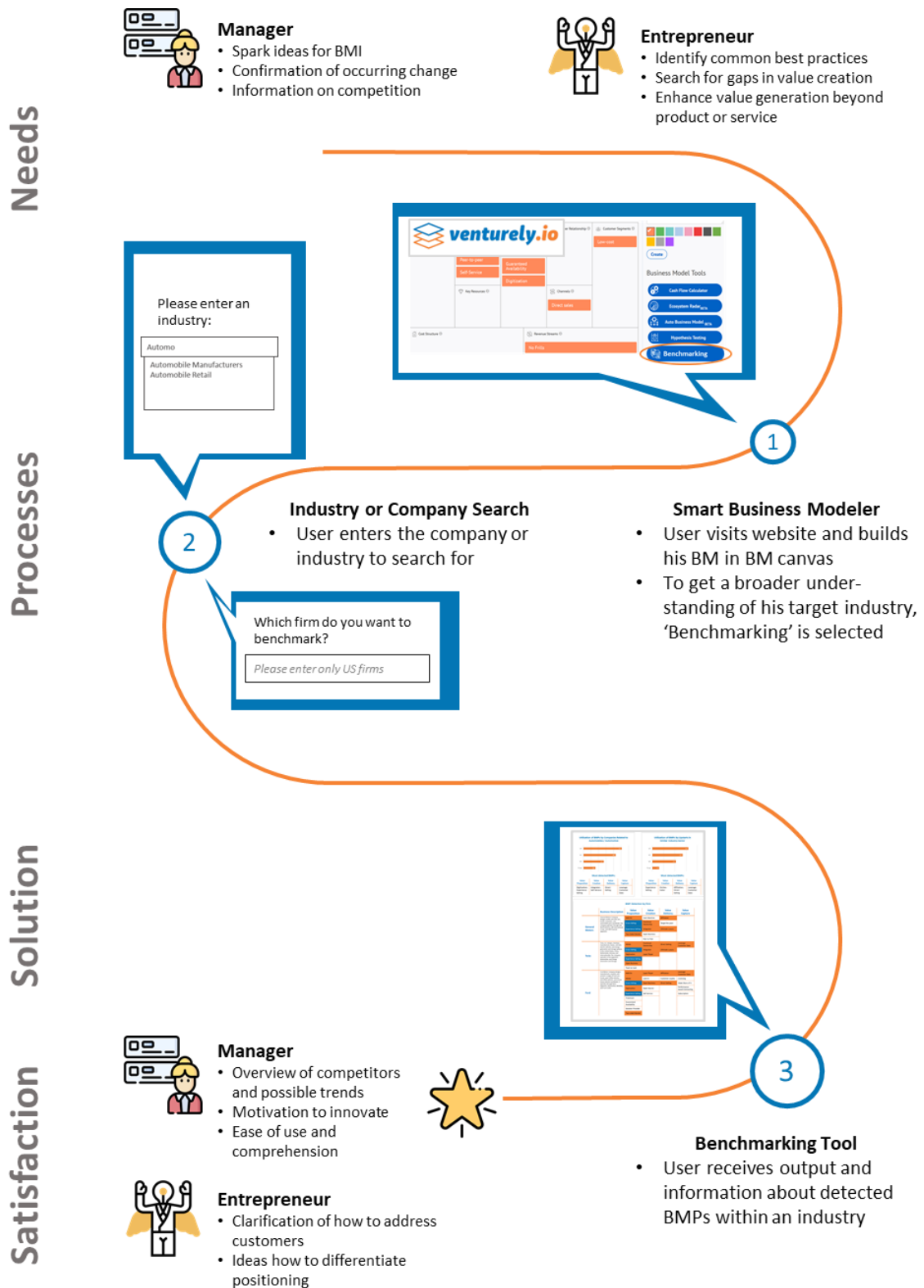
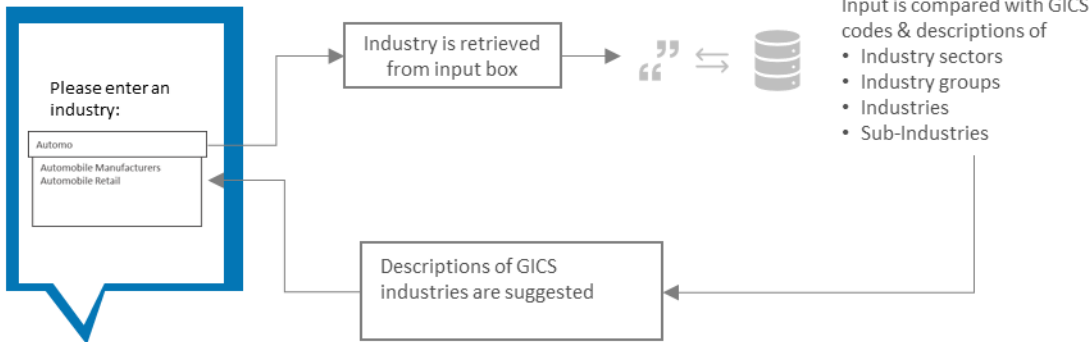


Figure 15: Customer Journey of Smart Business Modeling Tool

As previously described, the benchmarking tool will be implemented into a platform for smart business modelling and thereby complement further tools. Hereby, back-end processes have been discussed with a ML expert and are illustrated in . The implementation will continue after the conclusion of present thesis.

Back-end Processes

2 Industry Search – Users must choose from a suggestion table of GICS industries



Company Search – Users must choose from a suggestion list of companies

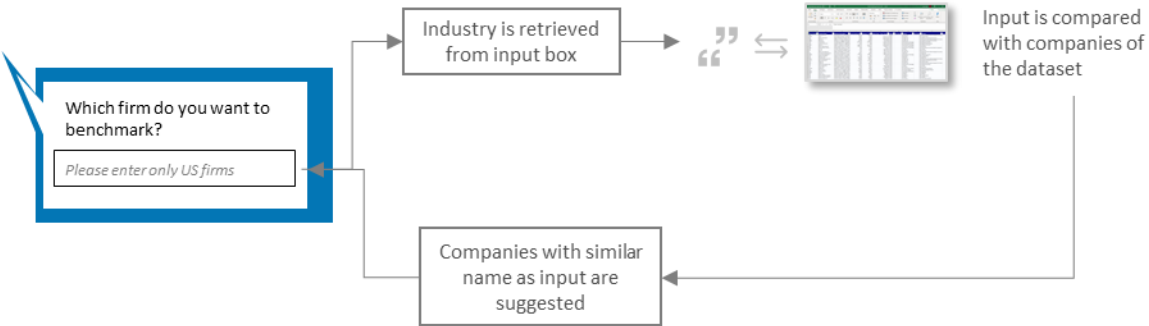


Figure 16: Back-End Processes of Business Modeling Tool