Master Thesis

University of Twente
Technische Universität Berlin

M.Sc. Business Administration
M.Sc. Innovation Management and Entrepreneurship

Corporate Foresight: The identification of potentially disruptive innovations from startups

A methodological framework to utilize startups as a source of corporate innovations

University of Twente
First supervisor: Dr. M.L. (Michel) Ehrenhard
Second supervisor: Prof. Dr. A.B.J.M. (Fons) Wijnhoven

Technische Universität Berlin
First supervisor: Prof. Dr.-Ing. Robert Dust
Second supervisor: Dipl.-Ing. Matthias Trotz

Name: Philipp Macheleidt
Student No: s1749641/362772
E-Mail: p.macheleidt@outlook.de
Date: December 12th, 2016

Declaration of Authorship

I hereby declare that I have written this Master's Thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work, including tables, maps and figures, which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

Berlin, 12th December, 2016

Philipp Macheleidt

Abstract

Corporate foresight is concerned with the observation and interpretation of environmental change factors and the identification of potential innovation in order to lay the foundation for future competitive advantage. Startup firms are recognized as a hotspot for ideas and innovations, in particularly for ones of a radical or disruptive nature. However, startups are not yet considered as a source for environmental information and innovations within the corporate foresight literature. The aim of this thesis is to amplify and complement the corporate foresight process and adjacent research streams by extending the environmental scanning and the identification of opportunities and innovations into the startup context. By conducting a comprehensive literature review the different research perspectives or fields of academic contribution relevant for corporate foresight were outlined and the role or value of startups within each perspective was examined. Furthermore, a methodological framework to identify startups from a startup database utilizing a text mining technique was proposed and applied. The findings show that startups should be considered as a valuable source for environmental information and innovation within corporate foresight. Furthermore, the results imply that the proposed methodological framework is an effective approach for the identification of relevant startups for a particular topic by utilizing a startup database. Firstly, the thesis contributes to the research area of corporate foresight by outlining the academic value of startups. Secondly, it contributes by proposing an approach to discover relevant startups and thus addresses requests from the corporate foresight literature for a structured methodological approach on how environmental information and innovation can be identified.

Keywords: corporate foresight, startup identification, innovation, environmental scanning, text mining

Management Summary

The competitiveness of a firm determines its existence in the long term. Innovations have been identified as an important enabler to create and sustain competitive advantage. A constant monitoring of the corporate environment for changes enables the firm to stay ahead of the competition. The intention behind the monitoring is to identify possible discontinuities early, specifically disruptions or disruptive innovations and empower the company to react before they occur or at least before they turn into a threat. In order to develop, recognize or identify innovations it is necessary to observe and scan the corporate environment for indications of change and innovations. A large amount of innovations come from nascent entrepreneurs or startups, specifically ones of a radical or disruptive nature. These disruptive innovations pose a particular threat to incumbent firms as they possess the potential to disrupt the existing market.

Although startups are considered to pose a threat for incumbent firms because of their characteristics to discover change early and react to it accordingly, corporate foresight does not yet consider startups as a source for environmental information or innovations specifically. Furthermore, current foresight activities rely heavily on the contribution of experts and are necessitating manual effort for scanning and searching the corporate environment. This requires a considerable amount of resources and is a costly process. Moreover, several scholars of the field outline the need for new and improved tools and methods to sense environmental change and identify innovations within corporate foresight.

In order to address the outlined problems, it is suggested that incumbent firms should incorporate startups as a potential source for environmental information and innovation into the corporate foresight activities. Furthermore, it is suggested to take advantage of a not yet well recognized research stream within foresight which is concerned with specific methodological approaches for the identification of environmental information as well as the discovery of opportunities, technologies, innovations or ideas. Text mining as a technique to automate certain parts of the scanning and identification process is dominantly applied within that stream of the literature and was also utilized within the research project at hand. It is suggested to utilize a startup database to identify startups potentially relevant to the topic under investigation. In particular, it is proposed to employ an automatic extraction of keywords from professional journals and utilize these keywords, characteristic for the topic at hand to identify startups from the startup database.

The research contributes by outlining from a theoretical perspective why corporate foresight should consider startups as a source for environmental information as well as innovations and

that they should be incorporated more actively into the corporate foresight process. Furthermore, the research contributes by outlining a methodological approach for the identification of startups, allowing to discover startups within a particular topic. It also addresses the drawback of current foresight tools as it reduces the amount of required expert resources. Moreover, the research exemplarily shows the identification of startups in the context of the connected vehicle and delivers practitioners a startup map. The startup map delivers an overview of different areas related and important with regard to the connected vehicle and startups that are active within this area. Through the methodological approach the research also addresses the call within the corporate foresight literature for new, more efficient and advanced methods for the identification of environmental information and innovations.

Table of contents

List of Figures

List of Tables

List of Abbreviations

DBI	Davies-Bouldin index
ECU	electronic control unit
ICT	information and communication technology
IDF	inverse document frequency
IFTTT	If This Than That
OBD	onboard diagnostic
OEM	original equipment manufacturer
SaaS	software as a service
TF	term frequency
TF-IDF	term frequency - inverse document frequency
V2I	vehicle-to-internet
V2R	vehicle-to-road infrastructure
V2S	vehicle-to-sensor on-board
V2V	vehicle-to-vehicle

1 Introduction

1.1 Relevance of the topic

"Imagination is more important than knowledge. Knowledge is limited to all we now know and understand, while imagination embraces the entire world, and all there ever will be to know and understand."

Albert Einstein

The competitiveness of a corporation determines its competitive advantage and therefore its existence in the long term (Porter, 1990). An increasingly dynamic and competitive environment (Talay & Townsend, 2015) makes it crucially important to achieve competitive advantage (Sirmon, Hitt, Arregle, & Campbell, 2010). Furthermore, innovations have been identified as an important enabler for competitive advantage by a variety of scholars (McGrath & Ming-Hone Tsai, 1996; Porter, 1990; Schulze, MacDuffie, & Täube, 2015). It can be achieved through external and internal sources whereupon innovations are addressing the both (Baark, Antonio, Lo, & Sharif, 2011; Barney, 1995).

In order to continuously innovate and stay ahead of the competition, it is necessary to constantly monitor the corporate environment to react to changes (Rohrbeck & Bade, 2012; Ruff, 2015; Talay & Townsend, 2015). Thereby, the intention behind the monitoring of the company's environment is to identify possible discontinuities early, specifically disruptions or disruptive innovations and empower the company to react before they occur or at least before they turn into a threat for the company.

The concept of disruptive innovation was introduced by Christensen (1997/2011). He defined disruptive innovations as innovations which are underperforming when they are introduced but which have the potential to outpace existing solutions over time and even define new markets as well as attract new customers (Christensen, 1997/2011). Mainly nascent firms or entrepreneurs create and introduce these disruptive innovations (Reid, Roberts, & Moore, 2015; Weiblen & Chesbrough, 2015). Such entrepreneurs, also called startups, are considered as more innovative regarding radical innovations while incumbents perform better with incremental innovation (Alvarez & Barney, 2001; Baumol, 2004; Henkel, Rønde, & Wagner, 2015; Reid et al., 2015).

Moreover, historical examples show that incumbents struggle with disruptive innovations and subsequently lose economic ground or even exit the market. Companies that struggled with the adoption of disruptive innovation are for example Kodak, Nokia or as a more recent example the multinational energy utilities such as EON, RWE, ENBW, and Vattenfall (Lucas & Goh, 2009; O'Reilly & Tushman, 2013; Richter, 2013; Vecchiato, 2015).

The concept of corporate foresight addresses the problem of a constantly changing environment and the deprivation of competitive advantage and market position (Albright, 2004; Rohrbeck, Battistella, & Huizingh, 2015). It enhances the identification, observation and interpretation of corporate environmental changes and potential opportunities by determining possible implications as well as responses (Baskarada, Shrimpton, Ng, Cox, & Saritas, 2016; Day & Schoemaker, 2004a). Innovations as an important source of competitive advantage could be one implication or response (Day & Schoemaker, 2004a; Townsend & Calantone, 2014). Scholars have researched the observation, scanning and monitoring of a company's environment, the adaption to external changes and the concept of disruptive innovation extensively (Rohrbeck, Battistella et al., 2015; Yu & Hang, 2010).

Despite the intensive research about corporate foresight as well as adjacent research streams, startups and small entrepreneurial firms are not considered as potential sources for environmental information or innovations, at least not mentioned in the corporate foresight literature in particular (Durst, Durst, Kolonko, Neef, & Greif, 2015; Horton, 1999). However, they are recognized as a source of innovations (Zahra, Sapienza, & Davidsson, 2006) in the stream of corporate venturing and open innovation in theory and practice. A variety of concepts, such as corporate venturing, corporate incubation (Roessler & Velamuri, 2015), corporate acceleration (Kohler, 2016), venture capital (Chemmanur, Loutskina, & Tian, 2014) or venture acquisition (Henkel et al., 2015) can enhance corporate innovation. Such concepts are widely researched by academic scholars (Weiblen & Chesbrough, 2015). In practice, the development substantiates in several corporate initiatives to internalize startups and entrepreneurial innovations across industry borders (see Table 5, p.37).

Although the research stream on corporate venturing and open innovation covers startups as a potential source of innovation, it does not include how these startups might be identified, at least not in a structured and comprehensive manner. Pauwels, Clarysse, Wright, and van Hove (2016) outline two basic possibilities to identify startups, either through an open call and application process or via event based scouting activities. These identification approaches are disadvantageous for corporations by limiting the possibilities to identify startups that are most beneficial for the corporation as they limit the choice to those actively applying or participating. This drawback is recognized by researches of the field as well. Hathaway (2016) for example proposes to use investor or startup databases to identify accelerator programs in his research. Moreover, a conference proceeding that occurred just during the course of this thesis utilized a startup database to identify new digital business models in the mobility sector (Remane, Hildebrandt, Hanelt, & Kolbe, 2016). This elucidates that there is a need to utilize startup databases to identify startups for a variety of reasons.

In contrast, the rather unstructured research field, which is covering topics like opportunity discovery, technology intelligence or technology discovery, exhibit approaches and methodologies to discover weak signals (Thorleuchter & van den Poel, 2015; Yoon, 2012), environmental information (Seo et al., 2016), white spots (Yoon, Park, & Kim, 2013), technological opportunities (Yoon et al., 2015), or ideas (Thorleuchter, van den Poel, & Prinzie, 2010). However, the link between corporate foresight and the research field on opportunity, technology and innovation discovery is missing. Heuschneider and Herstatt (2016) support this line of argumentation as they emphasize that there is a gap in the literature about the detection of future trends and discontinuities in the corporate environment. Nevertheless, they do not make the link to the literature about technology and opportunity discovery. The linkage between both corporate foresight and opportunity, technology or innovation discovery could be the first step to close the existing gap in the field of corporate foresight. It would support the development of more practical approaches on how to identify relevant environmental information or innovation, as requested by Rohrbeck, Battistella et al. (2015), Heger and Rohrbeck (2012) and (Paliokaitė & Pačėsa, 2015). These methodological approaches and proposed techniques could be valuable within corporate foresight, although they do not cover the question how startups as an extremely important source for the internalization of innovations can be identified. The existing approaches focus mainly on patents as a source for information (Jeong & Yoon, 2015; Yoon, Park, & Coh, 2014; Zhang et al., 2016).

Therefore, this thesis will close the gap between the corporate foresight, corporate venturing and open innovation as well as the opportunity discovery literature by theoretically connecting all three research streams and by proposing a methodological approach on how startups can be actively identified by companies.

The research is conducted exemplarily within the area of connected vehicles. Stimulated by the increasing connectedness of everyday's life and due to the benefits regarding safety and functional scope provided by this technical development, the car will increasingly transform into an always online consumer product. Examples which already indicate the transformation of passenger cars are BMW's integration of the web-based service If This Than That (IFTTT) (BMW AG, 2016b), Hyundai's connected car roadmap (Hyundai, 2016) or Daimler's broad product portfolio on smartphone-enabled mobility services (Daimler AG, 2015). The connected car is seen as an evolutionary stage in the development of automobiles. It can be described as a car that collects, processes, interprets and utilizes all incoming and outgoing data flows and actively communicates with vehicles, infrastructure, and other technological devices (Johanning & Mildner, 2015).

1.2 Research goal and research questions

Based on the preliminary illustration of the theoretical underpinnings and the identified gap, this research aims to develop a framework to recognize and identify environmental information, opportunities and potential disruptive innovations from the startup context. Particularly, this thesis theoretically examines the viability of startups as a source for such information derived from the academic literature. Moreover, it proposes a method to identify relevant startups by utilizing data from a business information database. By doing that, the research streams of corporate foresight, corporate venturing as well as the research field on opportunity, idea, and innovation discovery will be connected.

The aim is to amplify and complement the corporate foresight process and adjacent research streams by extending the environmental scanning as well as opportunity recognition into the startup context. The thesis's ambition is to empower a derivation of environmental information and potential opportunities or innovations to enhance the ability of companies to be innovative and thereby create competitive advantage. The present thesis addresses the following general research question and the subsequent two sub-research questions:

- How can startups be utilized as a source for corporate innovations within corporate foresight?
  - What role do startups play as a potential source of corporate innovations?
  - How can startups be identified utilizing a startup database?

1.3 Contributions

The contribution of the thesis is twofold. On the one hand, the thesis contributes to theory and practice by theoretically closing the identified gap between the different research streams and outlines that there is a need to incorporate startups more actively into the corporate foresight activities. Moreover, it is shown that there exists a requirement for a more thorough methodological approach to identify startups. On the other hand, it contributes to theory and practice by proposing a methodological approach on how startups, a valuable source of innovation, can be identified from a startup database. A more detail description of the contributions can be found in Chapter 5.

1.4 Structure of the thesis

The research is structured into five chapters with the aim to address the identified gap in the literature and answer the respecting research questions. The general description of the topic as well as its relevance is outlined in Chapter 1. Chapter 2 provides a detailed review of the existing literature, addressing different perspectives of corporate foresight, and the role as well as contribution of startups within each perspective. Furthermore, the review investigates different sources for innovation and the role of startups as a source for innovations. The review also looks into a specific literature stream about methodological approaches of opportunity, technology or innovation discovery. Chapter 3 proposes, based on the theoretical examinations, a methodological approach on how startups can be identified from a startup database by utilizing text mining and similarity analysis. Chapter 4 presents the results of the research, showing an overview of startups that could be identified with the proposed method from the startup database. Chapter 5 summarizes the theoretical conclusions from the different corporate foresight perspectives as well as from the innovation and startup perspectives. Furthermore, it integrates the findings from the corporate foresight perspectives with the proposed methodological approach and the results of the research. Moreover, Chapter 5 outlines the contribution of the thesis from a theoretical as well as from a practical perspective, describes potential limitations and provides an outline for future research.

2 Theoretical framework

The theoretical framework will provide the methodological approach for the literature review insights into the main underlying theoretical concepts of this thesis as well as important definitions.

2.1 Methodological approach towards a comprehensive literature review

A literature review builds the foundation for a thesis and reviews prior relevant scientific work about the topic at hand (Webster & Watson, 2002). This literature review will be of an explanatory nature (Rowe, 2014), aiming to identify, acquire, extract and synthesize existing scientific knowledge (Tranfield, Denyer, Marcos, & Burr, 2004). A review of the existing scientific literature from academia and practice is considered an essential component of every thesis or dissertation (Booth, Sutton, & Papaioannou, 2016). By analyzing and synthesizing the theoretical work that has been undertaking prior, the researcher can examine *"[…] what we already know, what we need to know, and how we can get there […]"* (Schryen, 2013, p. 140). Furthermore, as described by Webster and Watson (2002), a thorough review is necessary to identify the gap in the literature.

This thesis conducted a literature review in order to ensure that it is explicit, transparent, objective, structured and reproducible (Booth et al., 2016). Moreover, it followed the description of a theoretical review as described by Paré, Trudel, Jaana, and Kitsiou (2015). Theoretical reviews draw *"[…] on existing conceptual and empirical studies to provide a context for identifying, describing, and transforming into a higher order of theoretical structure and various concepts, constructs or relationships"* (Paré et al., 2015, p. 188).

The literature review followed the proposed structure of Wolfswinkel, Furtmueller, and Wilderom (2013). The concept of Wolfswinkel et al. (2013) is preferred over the one of Döring and Bortz (2016) or Booth et al. (2016). The reason is primarily based on complexity issues of the latter two approaches, as they provide detailed descriptions but not a short and precise approach like the one proposed by Wolfswinkel et al. (2013). Both Wolfswinkel et al. (2013) and Döring and Bortz (2016) mention one vital issue about the review process, that it is an iterative process. This is decisive for the quality of the literature review, as the aspired coverage can only be achieved through several iterations. The process includes the following five steps, as proposed by Wolfswinkel et al. (2013):

1. Define the criteria (include and exclude), the scope
2. Search the respective sources

3. Select the relevant literature
4. Analyze the content of each paper
5. Present and synthesize the knowledge of different sources

Firstly, the scope, the timeframe and the journal quality were defined. The time frame for the reviewed literature was set to the past ten years. Nevertheless, some older academic contributions were employed due to the relevance as well. Furthermore, A*, A or B publications were used primarily but due to a lack of availability of high ranked publications about certain topics, some publications with a lower rank were incorporated additionally. The ABDC (2016) and the ABS (2015) ranking lists were used simultaneously as a reference for the ranking of the publications due to their differences and opacity of each rankings' quality evaluation (Sangster, 2015).

Secondly, for the search of relevant literature several scientific databases were utilized to gain a broad understanding and comprehensive picture of the academic literature. For the search, the following search engines and databases were employed: Google Scholar, EBSCO, SCOPUS, TU Berlin PRIMO and Web of Science. However, Google Scholar was used as the primary search engine as it provides the most comprehensive coverage in comparison to the databases Web of Science and SCOPUS (Harzing & Alakangas, 2016). Academic journals or books comprised the main sources for the present research.

Thirdly, reviews were initially scanned to get an overview of the current literature. Based on the acquired knowledge the ongoing search specifically looked for material on the search strings presented in Table 1. Table 1 also shows the number of identified literature per subject. The literature was prejudged by headllines and abstracts. If it was considered as appropriate the whole paper was reviewed. Moreover, forward and backward citation was checked to identify additionally relevant literature, as recommended by Wolfswinkel et al. (2013). Fourthly, the literature analysis was conducted similar to the process described by Wolfswinkel et al. (2013), and based on the grounded theory principles. Thus only the best available literature was used, scanned, and finally excerpts were extracted and combined (Wolfswinkel et al., 2013). Fifthly and lastly the knowledge was synthesized and supported by visualisations or tables. At the end of each chapter a short conclusion was drawn to outline the topics contribution to corporate foresight and/or the relevance for startups as a source for information within the corporate foresight perspective or stream of the literature.

Number of identified literature	
Corporate foresight	32
Strategic management perspective	81
Innovation management perspective	71
Future research perspective	24
Dynamic capability perspective	17
Causation and effectuation perspective	20
Corporate entrepreneurship perspective	44
Development of foresight	11
Methodological approaches for opportunity, technology and innovation discovery	12
Innovation and Startup	
Innovation	75
Startup	48
Connected vehicle	24

Table 1: Identified literature per topic

2.2 Corporate foresight

2.2.1 Definition of corporate foresight

"Corporate foresight permits an organization to lay the foundation for future competitive advantage. Corporate Foresight is identifying, observing and interpreting factors that induce change, determining possible organization-specific implications, and triggering appropriate organizational responses. Corporate foresight involves multiple stakeholders and creates value through providing access to critical resources ahead of competition, preparing the organization for change, and permitting the organization to steer proactively towards a desired future" (Rohrbeck, Battistella et al., 2015, p. 2).

The origins of foresight date back to the 1920s when it was recognized as being important for management decisions and firm performance (see chapter 2.2.8) (Amsteus, 2012). Today foresight is considered a process to systematically look into the future of science, technology, the economy and the society, to manage uncertainty, identify the most promising research areas and react accordingly (Martin, 1995; Vecchiato & Roveda, 2010). Furthermore, foresight was recognized as a valuable tool to individual and organizational learning by identifying possible and desirable futures (Baskarada et al., 2016). Figure 1 shows the general corporate foresight process.

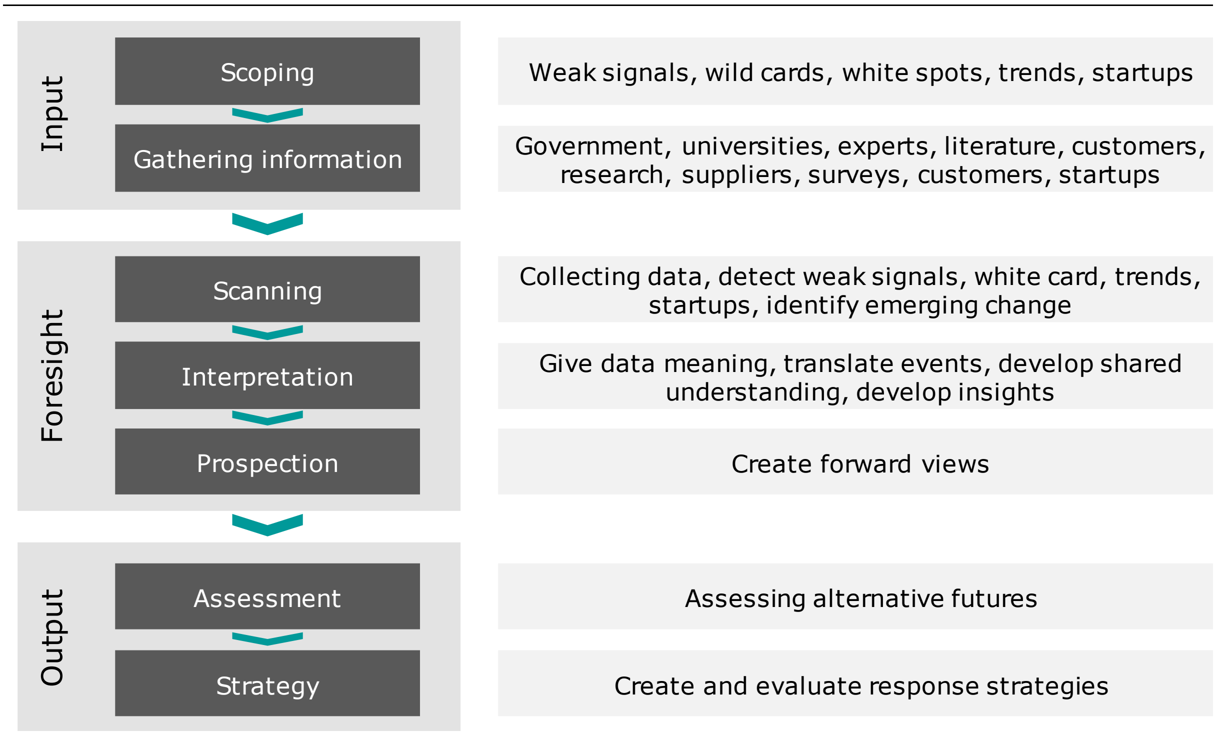

Figure 1: Corporate foresight process (Daft & Weick, 1984, p. 286; Durst et al., 2015, p. 93; Horton, 1999, p. 6; Kaserer, 2015, p. 61; Rohrbeck, Thom, & Arnold, 2015, p. 116)

Corporate foresight serves as an umbrella term for different foresight streams, foremost strategic foresight, innovation management, and future research (Rohrbeck & Bade, 2012; Rohrbeck & Gemünden, 2011).

Based on this origination from different research streams the previous academic literature has viewed corporate foresight from different perspectives (Müller, 2008; Rohrbeck, 2010a; Tyssen, 2012). This research will outline these different perspectives in the following and put an emphasis on innovation and the identification as of the topic of the thesis. Specifically, the following perspectives will be outline, strategic management perspective, the innovation management perspective, future research perspective, dynamic capability perspective, causation and effectuation perspective and corporate entrepreneurship perspective. Furthermore, a new and rather unstructured stream about methodological approaches for opportunity, technology and innovation discovery as well as the developments of corporate foresight will be examined.

2.2.2 Strategic management perspective

The strategic management perspective is of relevance within corporate foresight as major research stream that deals with achieving and sustaining competitive advantage (Barney, 1991) and subsequently with innovation and the creation of innovation (Porter, 1990). The major concern in strategy management and a prerequisite to achieving competitive advantage is the scanning of the corporate environment (Elenkov, 1997). To collect corporate environmental information two prominent concepts have been established in the academic literature. These concepts are referred to as environmental scanning (Ansoff, 1975) and peripheral vision (Day & Schoemaker, 2004a; Haeckel, 2004; Winter, 2004). Furthermore, scholars identified several types of environmental information a corporate entity should look at, such as weak signals (Ansoff, 1975), wild cards (Rockfellow, 1994), early warnings (Nikander & Eloranta, 1997) or white spots (Hammoud & Nash, 2014). In the following, all three strategic management systems as well as the possibilities to perceive the corporate environment will be addressed in more detail.

Porter (1996) outlines threes principles for strategy. Strategy has to create a unique value proposition (1), decide and make tradeoffs about competitive activities (2) and create fit among company activities (3) while continuously scanning the market and compare to competitors and customer needs to be flexible and able to respond. The origin of foresights' strategic management perspective goes back to a variety of concepts (see Table 2), mainly long range planning (Ansoff, Avner, Brandenburg, Portner, & Radosevich, 1970), strategic planning (Mason, 1969), and strategic management/strategic issue management (Ansoff, 1980).

All three theoretical strategic management concepts, long range planning, strategic planning and strategic issue management deal with the future of a firm, decisions to make at present and their impact on the future as well as information that are necessary to make these decisions in a profound manner (Drucker, 1959; Dutton & Ottensmeyer, 1987; Mintzberg, 1994). Long range planning is defined as *"[...] the continuous process of making present entrepreneurial (risk taking) decisions systematically and with the best possible knowledge of their futurity"* (Drucker, 1959, p. 240).

	Long-range planning	**Strategic planning**	**Strategic issue/surprise management**
Purpose	Anticipate growth and manage complexity	Change strategic thrusts	Prevent strategic surprises and respond to threats/opportunities
Basic assumptions	Past trends continue into the future	New trends and discontinuities	Discontinuities faster than response
Limiting assumption	The future will be like the past	Past strength apply to future thrusts. Strategic change in welcome	Future trends are OK

Table 2: Strategic management systems (Ansoff, 1980, p. 132)

Strategic planning, quite similarly is referred to as making upfront decisions about an unknown future state that is characterized by interdependent decisions about one or more possible futures that require a particular action to come true (Ackoff, 1970). Ansoff (1975) further outlines that many firms are lacking environmental information and that this lack of information about the companies surrounding can lead to strategic surprises, characterizing sudden, urgent and unfamiliar changes in the corporate environment, which are posing a potential threat or opportunity to the firm.

Furthermore, Ansoff (1980) created the concept of strategic issue management which gives firms the opportunity to have a profound strategic planning while being able to react quickly to environmental changes. This is necessary as the planning of the company's strategy is usually of a longer term nature and should not constantly be revised (Ansoff, 1980).

How to gain and select information about these environmental changes to enable and support foresight and planning is addressed in the academic literature with the two research streams of environmental scanning as well as peripheral vision.

Environmental scanning is described as a systematic approach to collect environmental information to inform management early about critical and important changes in the company's environment (Aguilar, 1967). Ansoff (1975) also contributed to the theory of environmental scanning, with his work about strategic surprises and weak signals. These weak signals are environmental information that seem to be random but become nonrandom, valuable information if the signals or dots are being connected via a framework or dominant logic (Haeckel, 2004). They allow the company to sense possible strategic surprises early, in the best case before they occur (Ansoff, 1975).

According to Hambrick (1982), scanning enables a company to learn about the surrounding and be able to react to changes in the environment. These changes refer to social, political, economic or technological changes (Jain, 1984). Furthermore, it is described as important for the detection of new or upcoming business practices in the corporate environment in order to avoid falling behind the competition (Albright, 2004). Through these scanning activities, an organization can respond to certain environmental changes and factors strategically (Albright, 2004). Not only is it possible to respond to changes, but it also increases firm performance due to a better awareness of possible changes and certain events (Daft, Parks, & Sormunen, 1988).

A second research stream that addresses the information retrieval from changes in the corporate environment is the research on peripheral vision. Day and Schoemaker (2004a) explain the concept of peripheral vision with an anthropomorphic analogy, as something that is at the outside edges of our central vision or central attention and we might miss out seeing. By neglecting these outside edges, they could quickly turn into something that opposes a huge risk or threat to us or in a transferred sense to the organization (Day & Schoemaker, 2004a).

Therefore, it is necessary to look actively into the periphery. According to Pina e Cunha and Chia (2007) peripheral vision involves *„[…] a cultivated sensitivity to the marginal, the hidden, the obscured and to what lies outside the frame of conscious attention"* (p.561). Important to note is that every time the focus changes, the periphery changes as well (Day & Schoemaker, 2004b). Peripheral vision enables a company to escape the dominant logic of the organization and to look beyond the obvious by overcoming the blinder of the periphery (Prahalad, 2004). By actively looking into the periphery, business opportunities, threats, and strategic blunders can be identified and shifts in the environment can be anticipated (Day & Schoemaker, 2004a). As mentioned earlier weak signals are one possibility to sense changes in the environment (Haeckel, 2004), however there are other concepts and categories like wildcards or white spots which can be a valuable source of environmental information as well (Dasgupta & Sanyal, 2009; Mendonça, Pina e Cunha, Kaivo-oja, & Ruff, 2004). Weak signals are referred to as a possible source of environmental information within the concepts of environmental scanning or peripheral vision (Ansoff, 1975; Haeckel, 2004). In the literature weak signals, early warning signs, signs, emerging issues, or future signs are used synonymously (Hiltunen, 2006, 2008). Ansoff (1975) defined weak signals as information or indications that are not strong enough for a profound strategic planning, but they might point a certain development towards a particular direction. Important to note about weak signals is that they are mostly recognized by pioneers and not by domain experts (Hiltunen, 2008).

Besides weak signals, there exists the concept of wild cards that also gained large scholarly interest (Cornish, 2003; Petersen, 1997; Rockfellow, 1994). *"A wild card is a description of an occurrence that is assumed to be improbable, but which would have large and immediate consequences for organizational stakeholders if it were to take place"* (Mendonça et al., 2004, p. 202). Although wild cards and weak signals are sometimes used synonymously, they are clearly separable from one another (Hiltunen, 2006). Weak signals are seemingly insignificant developments in the present, but they indicate probable events of the future, such as strategic surprises (Hiltunen, 2006; Roessler & Velamuri, 2015). Wild cards, on the other hand, also referred to under the term black swan are currently ongoing or past events that are unexpected, surprising events with enormous consequences (Hiltunen, 2006; Taleb, 2010).

A third important concept in the academic literature is the concept of white spots. White spots, also referred to as white spaces or opportunity spaces, describe product or service needs in the market which are currently not addressed (Hammoud & Nash, 2014; Rohrbeck, 2010b). In that sense, white spots are the identification of existing cavities in a company's portfolio, resulting in the possibility of detecting new business opportunities (Hammoud & Nash, 2014; Rohrbeck, 2010b).

Conclusion 1: The strategic management systems, environmental scanning as well as different sources for environmental information build the foundation of the corporate foresight research. Predominantly important within this research stream is the scanning of the corporate environment for information about change. Thereby pioneers, specialized groups and thus startups are considered a relevant source for information.

2.2.3 Innovation management perspective

An innovation is defined as:

> *"[…] the production or adoption, assimilation, and exploitation of a value-added novelty in economic and social spheres; renewal and enlargement of products, services, and markets; development of new methods of production; and establishment of new management systems. It is both a process and an outcome"* (Crossan & Apaydin, 2010, p. 1155).

A more detailed description of innovation can be found in chapter 2.3.1. Subsequently, innovation management *"[…] encompasses all the key areas that need to be mastered to develop successful products and services, efficiently and continuously"* (Hidalgo & Albors, 2008, p. 116). Innovation and thereof a successful innovation management has become an important factor for economic growth and firm performance (Demirel & Mazzucato, 2012; Sorescu, Chandy, & Prabhu, 2003; Tellis, Prabhu, & Chandy, 2009).

Innovation management can be considered a process that utilizes knowledge, technologies, and processes to generate innovations, such as new products, processes, technologies or business models (Galanakis, 2006). In general, the process is about generating knowledge, transforming that knowledge into new products, and ideally be successful with those products in the market (Galanakis, 2006). Figure 2 shows a synthesized process framework of innovation management.

Innovation management is pursued in the internal as well as the external environment of a corporation (Ortt & van der Duin, Patrick A., 2008). The firm's strategic planning, organizational culture, structure, and capabilities are an important internal environmental factor and contribute decisively to the innovation outcome (Ortt & van der Duin, Patrick A., 2008). External factors are on the other hand mainly referred to as sources of knowledge, inter-organizational relationships or the company's network (Cassiman & Veugelers, 2006; Damanpour, 1991; Ritter & Gemünden, 2003). The internal as well as the external factors will be discussed in more detail in the following.

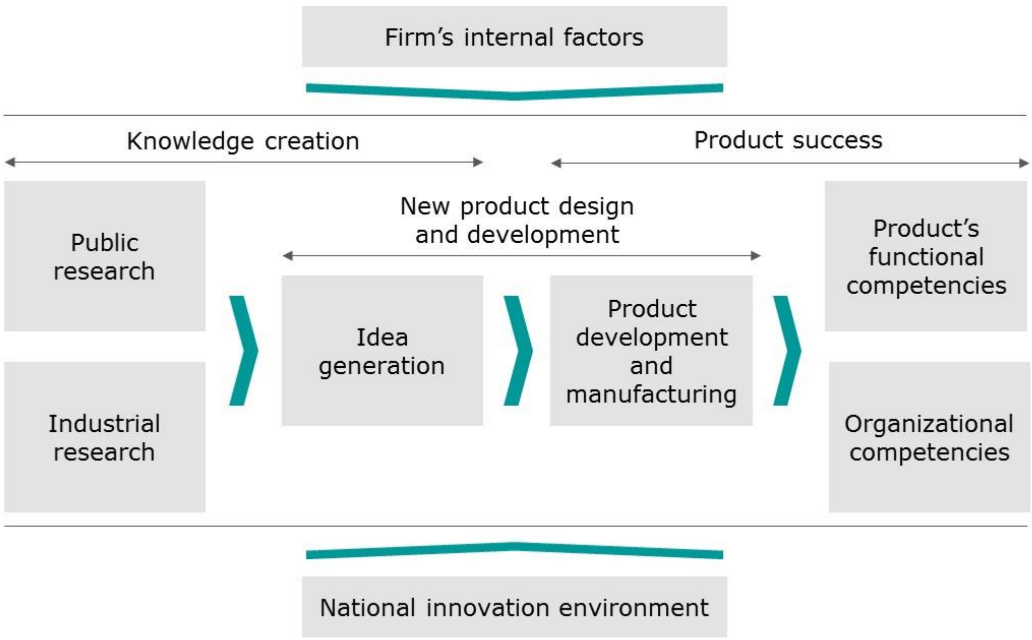

Figure 2: Innovation Process. Adapted from (Galanakis, 2006, p. 1231)

Organizational culture

The organizational culture deals with six internal environmental factors, that are the firm’s willingness to cannibalize, future market orientation, risk tolerance, use of incentives, product champions, and internal markets (Kostoff, Boylan, & Simons, 2004; Tellis et al., 2009). Willingness to cannibalize does refer to the firm’s willingness to potentially destroy the value of its prior investments. Future market orientation describes the extent the firm aligns its innovation efforts with the needs of potential future markets. Risk tolerance is indicating how much risk the company might take to reach the goal. Incentives display the reward system for innovation. Product champions align with the incentives and describe the extent a company promotes people who are highly innovative. Lastly, internal markets illustrate the internal autonomy and competition that exists. As startups, do often not possess a current market, it is easier for them to orient towards future markets as there is no need to cannibalize (Tellis et al., 2009).

Organizational structure

Organizational structure or dichotomous structure is considered as the distinction between exploration and exploitation (Beldow, Frese, Anderson, Erez, & Farr, 2009; March, 1991). Exploration is characterized by the terms search, variation, risk taking, experimentation, play, flexibility, discovery or innovation, while exploitation refers to refinement, choice, production, efficiency, selection, implementation or execution (March, 1991). In the context of an organizational structure, it is referred to exploration and exploitation as an ambidextrous organization (Tushman & O'Reilly, Charles A., III, 1996). This describes an organization which can develop new products and services on the one hand (exploration) and compete in existing markets by increasing efficiency and lower costs on the contrary (exploitation) (Tushman & O'Reilly, Charles A., III, 1996). Ambidexterity has positive effects on firm performance, growth (Geerts, Blindenbach-Driessen, & Gemmel, 2010), innovation (Tushman, Smith, Wood, Westerman, & O'Reilly, 2010), and firm survival (Hill & Birkinshaw, 2014).

Academic research distinguishes between three organizational structures to achieve ambidexterity, a sequential, a simultaneous and a contextual fashion (O'Reilly & Tushman, 2013). Sequential ambidexterity *"[…] reflects the temporal sequence of routines that balance exploration and exploitation in two successive time periods"* (Venkatraman, Lee, & Iyer, 2007, p. 8). Simultaneous ambidexterity describes the balancing of exploration and exploitation in the same time period (Venkatraman et al., 2007). This can, for example, be achieved via a subunit for either of the two tasks (O'Reilly & Tushman, 2013). Contextual ambidexterity argues that it is possible to pursue exploitation and exploration within one business unit by encouraging employees to decide on their own whether to gear towards alignment oriented activities (exploitation) or towards adaptability oriented activities (exploration) (Gibson & Birkinshaw, 2004). Other scholars, such as Raisch, Birkinshaw, Probst, and Tushman (2009) looked into the balancing of exploration and exploitation from a different perspective. They distinguished between differentiation or integration, individual or organizational, static or dynamic and internal vs. external balancing possibilities.

Stettner and Lavie (2014) provide further insight on the impact that exploration or exploitation activities via particular modes of operation, such as internal organization, alliances or acquisitions have. Their findings suggest that it is more beneficial to balance exploration and exploitation across modes than within modes, specifically that it is more advantageous to utilize internally oriented modes for exploitation and externally oriented modes (e.g. startups) for exploration. Furthermore, it is of particular benefit to explore in the acquisition model while exploiting via internally oriented modes (Stettner & Lavie, 2014).

With capabilities, as Teece (2007) noted, it is about a firm's dynamic capabilities, sensing, seizing and transforming to enable the exploration and exploitation of innovation opportunities. Under the view of corporate foresight, it is about these dynamic capabilities to discover and acquire new technologies as well as products and business models, link those emerging innovations to the needs of potential customers, initiate the appropriate and required R&D projects, find the right personnel and promote the required skills as well as enable the organization to develop radical and incremental innovations (Rohrbeck & Gemünden, 2011). Therefore, it is not only about exploring and exploiting innovations and opportunities but also about the exploration and exploitation of the capabilities that enable the firm to develop innovations or improve existing products and services (see Chapter 2.2.5) (Wang, Senaratne, & Rafiq, 2015).

Knowledge management

Darroch (2005) showed that the active management of the firm's knowledge resource has a significantly positive impact on innovation and partially on firm performance. Zahra and George (2002), Wallin and Krogh (2010) as well as Zhou and Wu (2009) furthermore conceptualized that when it comes to innovation, it is about the company's ability of successfully utilize that knowledge, thus about a firm's absorption capabilities. Absorption in this context refers to the acquisition, the assimilation, the transformation and the exploitation of knowledge (Zahra & George, 2002). Absorptive capacity and innovation show an inverted U-Shape relationship (Díaz-Díaz & Saá Pérez, 2014). While external knowledge is absolutely necessary to innovate, firms must be aware of the fact that in order utilize that knowledge to the greatest benefit there is a need to have previous internal knowledge in that particular area (Spithoven, Clarysse, & Knockaert, 2010). Sources of knowledge for innovation can be internal knowledge sources, competitors, customers, consulting firms, suppliers, universities, research institutes, patents, publications, databases, conferences, exhibitions, business clubs, investment networks, incubators or science parks (Leiponen & Helfat, 2009; Pittaway, Robertson, Munir, Denyer, & Neely, 2004). Leiponen and Helfat (2009) further investigated that the greater the breadth of knowledge sources, the higher the innovation success. The external knowledge is thereby sourced through three main sourcing modes, via information transfer from informal networks, R&D collaboration and technology acquisition (Kang & Kang, 2009). Startups could serve as a specific source that further broadens the sources of knowledge.

Nevertheless, it is not only about the knowledge and the absorption of knowledge within the organization, but also about the number of ideas that are generated (internal or external) and that have the possibility to influence the outcome towards success or failure significantly (Adams, Bessant, & Phelps, 2006).

Network

Knowledge, and especially the sources of knowledge are closely related to networks, as the network position can be beneficial for the company's innovativeness (Jenssen & NYBAKK, 2013). Phelps, Heidl, and Wadhwa (2012) identified three main components of a network: the knowledge network elements, the knowledge outcomes and the level of analysis. The elements include the properties network structure, relational nodes and knowledge flows. The outcomes refer to the creation, transfer, as well as adoption whereas the level of analysis, describes whether it is on an intrapersonal, intra-organizational or inter-organizational level (Phelps et al., 2012). Dagnino, Levanti, Minà, and Picone (2015) take a different approach and examine the knowledge transfer and opportunity for innovation within inter-organizational networks. With their bibliometric coupling analysis, they identified six main areas of inter-organizational network research: networks as a framework that sustain firm innovativeness in specific contexts (1), network dimensions and knowledge processes (2), networks as a means to access and share resources/knowledge (3), the interplay between firm and network characteristics and its effects on innovative processes (4), empirical research on networks in highly dynamic industries (5), the influence of industry knowledge domain's peculiarities on network dimensions and characteristics (6) (Dagnino et al., 2015).

Conclusion 2: The innovation management perspective contributes to corporate foresight by outlining the process and requirements for a successful innovation management. Specifically, how innovative ideas occur and how they can be fostered. Concerning startups, it shows that they possess particular benefits regarding organizational culture, organizational structure as well as knowledge.

2.2.4 Future(s) research perspective

Futures research perspective also referred to as future studies (Slaughter, 1996), foresight (Cuhls, 2003) or futurology (Sardar, 2010) engages in the topic on how corporations can prepare for upcoming environmental changes and developments (Eppink, 1981).
Van der Duin, Heger and Schlesinger (2014) identified four phases of development within futures research and linked it to the development of innovation management (see Table 3).

	Innovation process	**Futures research**
Phase 1 (1950s-1960s)	Technology push	Technology forecasting
Phase 2 (1960s-1970s)	Market pull	Technology assessment
Phase 3 (1970s-1980s)	Coupled innovation process	Exploratory futures research
Phase 4 (1980s-present)	Innovation in systems or networks	Networked foresight

Table 3: Phases of innovation management and future research. Adapted from (van der Duin, 2006, p. 275; van der Duin et al., 2014, p. 64)

Futures research emerged during the 1950s out of national foresight programs to derive necessary information to formulate national science and technology policies, as technology and innovation are considered as the drivers of economic growth (Blind, Cuhls, & Grupp, 1999; van der Duin, 2006). During the 1960s the utilization of research about the future entered the cor-

porate world, with Royal Dutch Shell as a prominent example of a company utilizing the scenario planning methodology to guide its strategic planning activities (Blind et al., 1999; Wack, 1985). During the 1970s future research functioned merely as a guiding tool for corporate vision (van der Duin, 2006). From the 1970s onwards future research transformed from trying to predict events to the exploration of possible future developments (McMaster, 1996; Rohrbeck, 2010a). Therefore, the research stream is nowadays characterized by the predominant logic, that it is not about the prediction of one future lying ahead but rather the identification of possible futures and future developments (Gordon, 1992; van der Duin, 2014; Wissema, 1981).
Wissema (1981) defines future research as a process that:

> *"[…] is concerned with the use of presently available concrete data to make statements about possible concrete developments in support of long-term policy forming and decision making of target groups (concerns, societies, political parties, institutions, and the like)" (p.29).*

To make suggestions about possible developments in the future, one key ingredient is creative imagination (Kuusi, Cuhls, & Steinmüller, 2015). It is seen as a key component as decisions, forecasts, and predictions cannot be based on facts, as there are no facts about the future. But future is imaginable, and by envisioning it, it might turn into reality, as it exists as a social construct in people's minds (van der Duin, 2014). However, some researchers argue, because there are no future facts, futures research is not a science as there is no science in the absence of facts (Kuusi et al., 2015). Van der Duin (2014) argues that it is less about the future itself, but about the applied methods and factors that lead to a successful futures research outcome. Future research needs to guarantee that the methods, process and subsequently the results ensure reliability and credibility, as proposed by Piirainen, Gonzalez, and Bragge (2012).
Methods that are prominently applied in futures research to enable a systematic exploration of future developments are for example Delphi studies, scenario analysis, roadmapping, backcasting, or mathematical models (Cook, Inayatullah, Burgman, Sutherland, & Wintle, 2014; van der Duin, 2014). Over time a variety of methods suitable for futures research emerged, a compiled list of applicable methods is provided by the Technology Futures Analysis Methods Working Group (2004). The application of these methods can thereby serve the objective to test the strategy, explore new business fields or new policy issues (van der Duin, 2014).
As van der Duin (2014) noted, the applied methods are dependent on the context and the surrounding conditions they are applied in, what turned out to be successful for one firm might not be successful for another firm. This is where the contingency approach comes into play, as the success depends on the organization's ability to adjust to the environmental conditions (Donaldson, 2001, 2006).

Conclusion 3: The futures research perspective is relevant for corporate foresight as it provides the methods that are necessary for its conduction. The methods do thereby vary, as there are no universal success factors and the methods need to be chosen dependent on the respecting environmental conditions.

2.2.5 Dynamic capability perspective

The dynamic capability perspective is based on two different concepts, the contingency theory, and the resource-based view. Within contingency theory, it is argued that there is not one single organizational structure that fits all organizations but that there are several factors, the contingency factors, which determine the most beneficial structure (Donaldson, 1999). The contingency factors can be for example the strategy, size or task uncertainty and are dependent on the company's environment (Donaldson, 1999). The goal, which is gaining a competitive advantage over rival firms, can be achieved by attaining the best fit between contingency factors and organizational structure (Donaldson, 2006). Within the resource-based view, on the other hand, it is argued that it is about a firm's resources, including physical-, human-, and organizational capital resources which have to be valuable (V), rare (R), imperfectly imitable (I) and non-substitutable (N) or VRIN to be able to achieve and sustain competitive advantage (Barney, 1986; Barney, 1991).

However, both concepts are based on static assumptions (Donaldson, 2006). The structural contingency theory rests on the assumption that firms can change to reach a better fit and thus higher performance, but once they changed they are in an equilibrium stage (Donaldson, 2006). The resource-based view is also considered static as the theory assumes that resources are distributed in a heterogeneous way across firms and that these resource allocations stay stable over time (Barney, 1991; Barney, 2001).

The dynamic capability perspective is now extending the contingency theory and the resource-based view and examines how to create and continuously refresh dynamic capabilities in dynamic markets (Ambrosini & Bowman, 2009; Blyler & Coff, 2003). In this respect, dynamic capabilities are defined as an approach to address dynamic environments by having the ability to integrate, build and reconfigure internal as well as external competencies (Teece, Pisano, & Shuen, 1997). Later academic work concludes that a firm needs sensing, seizing and reconfiguring capabilities to transform the existing capabilities and create new organizational and strategic routines to gain as well as maintain a competitive advantage (Eisenhardt & Martin, 2000; Teece, 2007).

Firms with strong dynamic capabilities require a strong entrepreneurial management and thus firms with strong dynamic capabilities are strongly entrepreneurial (Teece, 2007). Being entrepreneurial is of particular importance for sensing new capabilities (Teece, 2014).
According to Zahra et al. (2006), there are differences in the sensing, seizing and reconfiguring of dynamic capabilities between new ventures and established companies. This might have an impact on the fact that new ventures are exploring, creating and exploiting opportunities before established companies do so (Zahra et al., 2006). Furthermore, as Arend (2014) could show, young firms possess a slightly higher ability to develop appropriate dynamic capabilities and utilize them to affect firm performance positively.

Conclusion 4: The dynamic capability perspective contributes to the corporate foresight literature as corporate foresight can be regarded a dynamic capability as it enables a company to renew its resources in a dynamic market environment. Looking into startup context, it can be concluded that startups are valuable for incumbents as new ventures have different, expedient and superior skills in sensing, seizing and reconfiguring dynamic capabilities.

2.2.6 Causation and effectuation perspective

Causation and effectuation can be described as a distinct logic of human decision-making and action-taking in uncertain environments (Sarasvathy, 2001). Sarasvathy (2001) defines causation as a process that takes "*[…] a particular effect as given and focus[es] on selecting between means to create that effect*"(p. 245). Effectuation, on the other hand, is defined as a process that takes *"[…] a set of means as given and focus[es] on selecting between possible effects that can be created with that set of means*" (Sarasvathy, 2001, p. 245). More specifically, causation is a goal-directed approach where a predefined goal is sought to be achieved with the most efficient set of means, for example defining the destination for a trip and then select the best mode of transport (Berends, Jelinek, Reymen, & Stultiëns, 2014). Effectuation, on the other hand, takes a set of means as the starting point and aims, given those means, to create potential outcomes. Thus, planning a trip would examine potential modes of transport as well as resource constraints, such as time or budget and suggest potential destinations (Berends et al., 2014).
Furthermore, several scholars identified specific characteristics of causation and effectuation (Berends et al., 2014; Cai, Guo, Fei, & Liu, 2016; Chandler, DeTienne, McKelvie, & Mumford, 2011; Fisher, 2012; Werhahn, Mauer, Flatten, & Brettel, 2015). Chandler et al. (2011) outline in their research that causation is a unidimensional construct and effectuation a multidimen-

sional construct. Effectuation is considered a formative construct with experimentation, affordable loss and flexibility as independent sub-dimensions and pre-commitments as a shared dimension with causation (Chandler et al., 2011; Fisher, 2012; Werhahn et al., 2015). Under the first construct, experimentation, it is understood that multiple variations are developed to achieve commercialization, different ways to deliver and sell these products or services are developed and over the venture's development the product or service is changed substantially (Chandler et al., 2011; Fisher, 2012). The second construct, affordable loss, is considered as an entrepreneurial behavior that commits only limited amounts of resources to a venture at a time and examines how much loss is affordable (Fisher, 2012; Sarasvathy, 2001; Werhahn et al., 2015). Flexibility describes that the venture responds to upcoming opportunities, adapts the necessary resources accordingly and avoids actions that would restrict its flexibility (Fisher, 2012). Fourthly and lastly, pre-commitment characterizes that the entrepreneur establishes relationships with partners, suppliers, and customers in an early stage (Fisher, 2012; Sarasvathy, 2001). The second construct, causation, is considered one dimensional and characterized through the identification of an opportunity before developing anything towards it (Chandler et al., 2011).

Fisher (2012) as well as Reymen et al. (2015) showed that effectual and causational characteristics and decision-making procedures exist contemporarily in a venture and that effectuation specifically occurs in new ventures. Furthermore, as Read, Song, and Smit (2009) could show in their meta-analysis, effectual strategy, and decision making is significantly positive related to new venture performance. Similar results have been provided by Cai et al. (2016) who give evidence from Chinese new ventures and show a positive relationship between effectuation and performance.

The characteristics of small firms, such as low bureaucracy, rapid decision making, risk taking, motivated labor or effective internal communication (Vossen, 1998) enable the new venture to react to changing market demands or conditions faster (Chandy & Tellis, 2000). Reacting to these changing market conditions through innovations does often result in niche market activities or strategies (Fernhaber, McDougall, & Oviatt, 2007). However, these niche market innovations have, although they initially underperform compared to the mass market products, the potential to improve over time and become a mass market product that is superior to the existing product (Christensen, 1997/2011). Sood and Tellis (2011) describe it in more detail and show that formerly niche solutions have the potential to become disruptive for the existing mainstream market segment.

Another distinct characteristic of new ventures and small firms is that they often do not follow a traditional and well-structured new product development process, as large firms do (Scozzi, Garavelli, & Crowston, 2005). This again supports the argument that small and new ventures pursue approaches of effectuation rather than causation. However, as the venture grows and exists for a longer time period, the venture is more likely to change from effectual to causal logic (Reymen et al., 2015). By focusing on means and less on ends allows new ventures to be more open to new ideas and innovative approaches (Sarasvathy, Kumar, York, & Bhagavatula, 2014).

In summary, it can be said that the concepts of effectuation and causation has two distinct paths of entrepreneurial decision making, venture as well as idea creation show that entrepreneurs and the effectuation logic can serve as a decisive source for innovations. Furthermore, it is suggested that not only a firm's capabilities to predict the future but also the capability to create something new by adapting to the environment have to be improved (Chandler et al., 2011). Nevertheless, focusing on creating instead of predicting might be more beneficial with regard to innovation performance (Berends et al., 2014).

Conclusion 5: The causation and effectuation perspective contributes to corporate foresight as it describes two distinct logics of how opportunities and innovations are created. Concerning startups, it can be concluded that they tend to be more effectual than incumbents and thus show a higher likelihood of developing innovations specifically radical and disruptive innovations.

2.2.7 Corporate entrepreneurship perspective

Another field of research that is related to corporate foresight is the field of corporate entrepreneurship (Tyssen, 2012). As indicated in chapter 1.1 innovations are important for a corporation to achieve sustainable competitive advantage (Schulze et al., 2015). The entrepreneur is considered as the recognizing entity, creating or identifying these opportunities (Dyer, Gregersen, & Christensen, 2008). "*Entrepreneurship, in its narrowest sense, involves capturing ideas, converting them into products and, or services and then building a venture to take the product to market*" (Johnson, 2001, p. 138).

With regard to a corporation, entrepreneurship is referred to as the introduction, of new ideas, products, organizational structures, processes, or the creation of a new organization, either internal or external to the existing firm (Sharma & Chrisman, 1999) and known under the term corporate entrepreneurship (Zhao, 2005). Corporate entrepreneurship is also referred to in the literature as venture entrepreneurship (Tang & Koveos, 2004), innovation entrepreneurship (Tang & Koveos, 2004), intrapreneurship (Antoncic & Hisrich, 2001), corporate intrapreneurship (Dunlap-Hinkler, Kotabe, & Mudambi, 2010), and internal corporate venturing (Burgelman, 1983). The commonality between the concepts is that a corporation has to continuously reorientate itself (Urban & Wood, 2015) to renew (Phan, Wright, Ucbasaran, & Tan, 2009).

A variety of scholars has shown that this process of renewal consists of several underlying constructs. Guth and Ginsberg (1990) proposed that corporate entrepreneurship consists of two main domains, innovation and venturing within the organization on the one side and strategic renewal of established organizations on the other side. Kuratko and Audretsch (2013) do also share this separation in two domains and refer to these as corporate venturing and strategic entrepreneurship.

The domain of corporate venturing summarizes several methods to recognize, create or invest in business opportunities (Kuratko & Audretsch, 2013). Narayanan, Yang, and Zahra (2009) define it as a *"[...] set of organizational systems, processes and practices that focus on creating businesses in existing or new fields, markets or industries - using internal and external means" (p.59).* With internal corporate ventures the businesses or young ventures are typically created within the corporation, usually reside inside the corporation but may also reside outside the corporation as a semi-autonomous entity, like a spin-off (Kuratko & Audretsch, 2013; Kuratko, Hornsby, & Hayton, 2015; Phan et al., 2009). External corporate ventures on the other side are referring to businesses that are created outside the corporation and become part of the corporation through investment or acquisition (Kuratko et al., 2015; Kuratko & Audretsch, 2013; Phan et al., 2009).

The second domain, strategic entrepreneurship, refers to activities characterized by risk taking and aims to increase the corporation's competitive advantage through innovations (Phan et al., 2009). These innovations can be changes with regard to the firm’s strategy, product offerings, organizational structures, or business model (Kuratko & Audretsch, 2013). Furthermore, like Ireland, Hitt, and Sirmon (2003) note, strategic entrepreneurship involves opportunity-seeking and advantage-seeking behaviors simultaneously.

Within corporate entrepreneurship and thus within corporate innovation as well as innovation management, opportunities play an extremely important, if not the most important role (Stevenson & Jarillo, 1990; Urban & Wood, 2015). Entrepreneurship scholars developed several concepts of how opportunities become an innovation and thus contribute to competitive advantage. Opportunities can be searched (Ireland et al., 2003), spotted (Schildt, Zahra, & Sillanpaa, 2006; Stevenson & Jarillo, 1990), recognized (Baron, 2006; DeTienne & Chandler, 2004; Martín-de Castro & Fischer, 2011; O'Connor & Rice, 2001), identified (Corbett, 2005; Zahra, Korri, & Yu, 2005) or created (Alvarez & Barney, 2007; DeTienne & Chandler, 2004). No matter which concept, important is that the entrepreneur or firm needs to be aware of the opportunity (Zahra, 2015), or as Ardichvili, Cardozo, and Ray (2003) argue that the entrepreneur or firm has to notice factors within his field of expertise that are recognized and evaluated as a potential business opportunity.

Conclusion 6: The research stream of corporate entrepreneurship deals with the creation, recognition and identification of opportunities and their contribution to the corporate foresight literature by examining how a corporation can create innovations. Concerning startups, it can be concluded that the research stream of corporate entrepreneurship specifically points out that the opportunity creation can be internal or external to the company and that external ventures should be considered a valuable source for opportunities.

2.2.8 Development of corporate foresight

As mentioned earlier, foresight dates back to the 1920s (Amsteus, 2012). Rohrbeck, Battistella et al. (2015) identified four main development stages of corporate foresight. Around the 1950s, the concept of corporate foresight emerged from the French prospective school by Gaston Berger and from the foresight school by Herman Kahn (Rohrbeck, Battistella et al., 2015). During the 1960s and 1970s foresight experienced a more practical, development with the introduction of methodologies and tools to conduct foresight, most famously the scenario technique introduced by Royal Dutch Shell (Rohrbeck, Battistella et al., 2015). The methods and processes were further improved during the 1980s and 1990s and the role of corporate foresight from a rather forecast-oriented approach shifted towards an increased implementation of the corporate foresight process into the strategic decision making and innovation management (Rohrbeck, Battistella et al., 2015). This time period also created new methods relevant to foresight, most prominently the technology road-mapping approach (Bray & Garcia, 1997). Furthermore, during this period, corporate foresight shifted from a rather static approach to a more dynamic approach (Slaughter, 1990). The last phase, from the 2000s onwards, is mainly characterized by a growing integration of corporate foresight into the organization and the organizational routines, such as research & development, innovation management, strategic management, risk management and corporate development (Rohrbeck, Battistella et al., 2015).

In addition, Daheim and Uerz (2008) identified four waves of corporate foresight, (1) the expert-based wave, (2) the model-based wave, (3) the trend-based foresight, and (4) open foresight which are comparable to the stages identified by Rohrbeck, Battistella et al. (2015).

In addition to the last wave about an increasingly open and integrated corporate foresight recent scientific publication indicate a new stage or trend, the increasing support of corporate foresight procedures by information communication technology (ICT). Information and communication technology is used more and more to support corporate foresight activities (Rohrbeck, Thom et al., 2015). Keller and Gracht (2014) showed with a recent Delphi study that ICT support for corporate foresight activities will play an increasingly important role. They also indicate that ICT-applications will be used more and more for the sense-making of information as well as for deriving the right strategy and action instead of a sole collection of information (Keller & Gracht, 2014). Both, the gaining importance for ICT-based applications as well as the utilization for more complex tasks can be seen by looking into recent academic contributions in the particular field (Gracht, Bañuls, Turoff, Skulimowski, & Gordon, 2015; Keller & Gracht, 2014). Gracht et al. (2015) define these ICT-based applications and methods as foresight support systems, specifically referring to the term as:

> *"collaborative computer-based systems aimed at supporting (1) communication, (2) statistical and qualitative data analysis, including expert assessments (3) decision modeling (4) and rules of order in foresight processes"* (p. 2).

Rohrbeck, Thom et al. (2015) outlined in their research about the PEACOQ Scouting Tool how such an IT tool for foresight could look like. Furthermore, they identified several other applications that are used in foresight, such as news reader, internal libraries, tagging platforms, wikis or online idea competitions.

Conclusion 7: Recent corporate foresight developments in utilizing ICT-Tools to support and complement traditional corporate foresight activities advance more and more into an automated information retrieval and sense making. Future corporate foresight activities will most likely expand into this domain and use a variety of data sources to identify relevant environmental information, opportunities and innovation.

2.2.9 Methodological approaches for opportunity, technology and innovation discovery

Additionally, to the identified tool in the corporate foresight literature, mentioned in the previous chapter there are further academic contributions on specific methods and tools that are valuable within foresight (Kayser & Blind, 2016). However this research field is not yet connected to the corporate foresight literature although corporate foresight is in search for new foresight methodologies (Heuschneider & Herstatt, 2016; Paliokaitė & Pačėsa, 2015). Yoon (2012), for example, proposed a quantitative ICT-based method for weak signal detection by utilizing automatic keyword extraction through text mining. Thorleuchter and van den Poel (2015) took the idea of an automated weak signal identification one step further and proposed an approach for idea mining for web based weak signal detection. This approach is, compared to other ICT-based approaches superior and viable as it reflects the topics and trends in that particular area that have been identified in the academic literature (Thorleuchter & van den Poel, 2015). Yoon, Phaal, and Probert (2008) proposed to use text mining and a morphology analysis to identify potential product and technology opportunities. They further proposed to utilize product and technology information from product manuals and patent documents to append and complement the technology road mapping and thus integrate the ICT-application into traditional corporate foresight methods (Yoon et al., 2008). Zhang et al. (2016) further elaborated on that topic and proposed a mechanism based on an analytical method to cluster terms and phrases to identify relevant technological topics and finally transfer the information into a technology roadmap. Seo et al. (2016) proposed to utilize text mining techniques to identify product information from a patent database. Subsequently, the patent information was used to identify product opportunities and evaluate them, based on the firm's internal capabilities (Seo et al., 2016). A similar approach was suggested by Yoon, Seo, Coh, Song, and Lee (2016) who proposed a three-step approach to identify product opportunities based on patent data. They recommended untapped products and product areas for the target firm. These products were identified through natural language processing techniques as well as collaborative filtering. The most promising product opportunities were selected and displayed in a visual map (Yoon et al., 2016).

Conclusion 8: This rather new field of research about specific methodological approaches is not yet connected to the field of corporate foresight, although corporate foresight is requesting more advanced methodological approaches to enhance the corporate foresight performance. Furthermore, it shows that a variety of data sources can serve to identify potential opportunities, technologies or innovations and support the corporate foresight process.

2.3 Innovation and startup

2.3.1 Innovation

Definition of an innovation

One of the early contributors to the concept of innovation was Joseph Alois Schumpeter, who looked at innovation from an economic perspective. In the book "*Theorie der Wirtschaftlichen Entwicklung*" he defined innovation as creating value through new combinations (Schumpeter, 1912, 1934/2004). Baregheh, Rowley, and Sambrook (2009) analyzed and summarized more than 60 definitions of the term innovation in their research. Their elaborations showed that the term innovation is studied and defined by several other disciplines and perspectives, such as business and management, organization studies, innovation and entrepreneurship, technology science and engineering, as well as marketing (Baregheh et al., 2009). As a result of their content analysis innovation was defined as:

> "[…] *a multi-stage process whereby organizations, [groups of individuals or individuals] transform ideas into new/improved products, service or processes, in order to advance, compete and differentiate themselves successfully in their marketplace.*" (Baregheh et al., 2009, p. 1334), [addition] by (Hidalgo & Albors, 2008, p. 114)

Another definition, based on a comprehensive literature review, covering academic contributions from the last 27 years is the review by Crossan and Apaydin (2010). They defined innovation as:

> *"[…] production or adoption, assimilation, and exploitation of a value-added novelty in economic and social spheres; renewal and enlargement of products, services, and markets; development of new methods of production; and establishment of new management systems. It is both a process and an outcome" (Crossan & Apaydin, 2010, p. 1155).*

Both definitions give a first understanding of the term innovation. Although both are reviews of a large amount of literature but still display certain differences. Both definitions were considered relevant within this report. Nevertheless, due to the extensive usage of the term innovation in a variety of contexts and disciplines it is much more complex and cannot be summarized holistically by a single sentence. Therefore, the specific attributes of innovations will be outlined in more detail.

Baregheh et al. (2009) identified six characteristics of an innovation: (1) the nature of an innovation, (2) the type of an innovation, (3) the stage of an innovation, (4) its social context, (5) its means and (6) the underlying aim of an innovation.

The nature of an innovation (1) indicates something new or at least significantly improved (Baregheh et al., 2009; Oecd, 2005). New or significantly improved can thereby be seen in the context of the firm, the market or the world (Oecd, 2005). As types of innovations (2) the literature mentions a product, a process, a marketing method or a new organizational method (Oecd, 2005). Additionally to those four basic types of innovations scholars have identified several other innovation types, for example product, process or technological innovation (Dosi, 1982). Furthermore, recent scholarly contributions have shown that it is not only about innovations on the product and process level but about business model innovations (Chesbrough, 2007). The third characteristic, stages of innovation (3), describes the different stages an innovation passes, from the generation of the idea to its commercialization (Baregheh et al., 2009). The social context of an innovation (4) describes in detail who is involved in the innovation, like the organization, the firm, the customer, the employee or the developer (Baregheh et al., 2009). The fifth attribute, the means of an innovation (5), refers to the requirements necessary for an innovation. These requirements are inter alia the internal environment, the resources, the idea, and the market which all have an influence on the innovation outcome (Baregheh et al., 2009). An innovation’s aim (6) is to succeed, to compete or to differentiate against existing solutions (Baregheh et al., 2009).

One characteristic that has not been addressed by Baregheh et al. (2009) is the classification of the innovation regarding its innovativeness. Innovativeness describes the degree of newness of an innovation (Garcia & Calantone, 2002). More specifically it is the discontinuity an innovation can bring to the market and thus the potential shift in science, technology, industry or society (Garcia & Calantone, 2002). Within corporate foresight, it is necessary to achieve sustainable competitive advantage and the innovativeness of an innovation has a decisive impact on that (Barney, 1991; Markides, 2006; Porter, 1990). Several scholars have contributed to the question of how innovations can be classified in terms of their innovativeness. However, the literature does not provide one clear overview of the different classifications. Therefore academic contributions that try to differentiate between the various degrees of innovativeness were synthesized in this chapter into one comprehensive classification framework. Appendix 1 provides a detailed overview of different classifications and the respective definitions.

The existing literature on the classification of innovation differentiates between one or two dimensional scales (Abernathy & Clark, 1985; Assink, 2006; Chandy & Tellis, 1998; Christensen, 1997/2011; Dewar & Dutton, 1986; Porter, 1985; Robertson, 1967; Veryzer, 1998). However, there is no commonly applied framework in the scientific literature on how to classify different levels of innovativeness. Nevertheless, based on the definitions it can be said that there are ample overlaps in the underlying characteristic (compare Appendix 1). Incremental innovations, also referred to as continuous or conservative innovations can be defined as an alteration of an existing product with minor changes or adjustments to improve performance, increase the customers desire for the product and lower the cost while employing the same technologies without providing new benefits. (Assink, 2006; Chandy & Tellis, 1998; Dewar & Dutton, 1986; Norman & Verganti, 2014; Robertson, 1967; Veryzer, 1998). Radical innovations, also referred to as discontinuous or competence destroying innovations can be defined as a new product, process or behavioral pattern that leads to fundamental changes in the status quo, e.g. a substantially new technology and considerable greater customer benefits (Assink, 2006; Dewar & Dutton, 1986; Norman & Verganti, 2014). These changes are characterized by a discontinuity with the past as well as high instability and high uncertainty but have the potential to create an entirely new business practice (Assink, 2006; Corso & Pellegrini, 2007; Norman & Verganti, 2014).

Additionally, to the one-dimensional differentiation between incremental and radical innovations many authors propose a two-dimensional classification. Most authors use the distinction between incremental and radical innovations as a basis scale within the two-dimensional classification. Examples of a two-dimensional differentiation are the distinction between technological- and product capability (Veryzer, 1998), between technology/process/concept and market (Assink, 2006) or between technology and meaning (Norman & Verganti, 2014).

Additionally, Christensen (1997/2011) describes in his work the differentiation between sustaining and disruptive innovation. Sustaining innovations are defined as innovations improving *"[…] the performance of established products, along the dimensions of performance that mainstream customers in major markets have historically valued" (Christensen, 1997/2011, p. xviii).* Disruptive innovations on the other side

> *"[…] underperform established products in mainstream markets. But they have other features that a few fringe (and generally new) customers value. Products based on disruptive [innovations] are typically cheaper, simpler, smaller, and, frequently, more convenient to use"* (Christensen, 1997/2011, p. xviii).

Figure 3 illustrates the concept of disruptive innovations by showing the characteristics how to transform a niche product into a disruption for the mainstream market. At time T_1 mainstream customers value product characteristics of the dominant technology on the primary dimension while niche customers value the product features of the new technology on the secondary dimension (Sood & Tellis, 2011). Over time both technologies improve, but the new technology improves on the primary dimension and becomes a mainstream product itself, meeting or even overachieving the requirements of the mainstream customers (Sood & Tellis, 2011). Furthermore, Christensen (1997/2011) as well as Wan, Williamson, and Yin (2015) point out that sustaining and disruptive innovation can be of an incremental or radical nature.

Based on these analyzed two-dimensional concepts an integrated model is proposed (see Figure 4). The integrated model combines the framework of Assink (2006) with the model of Norman and Verganti (2014), with technology on one dimension and market on the second dimension. Furthermore, the scales on both dimensions consider innovations of an incremental and radical nature. Technical as well as market innovations can either be incremental or radical. However, a breakthrough innovation appears when the technological and the market side are of a radical nature. Additionally, to the two dimensions and the different scales, Christensen's concept of sustaining and disruptive innovations has been incorporated in the new model.

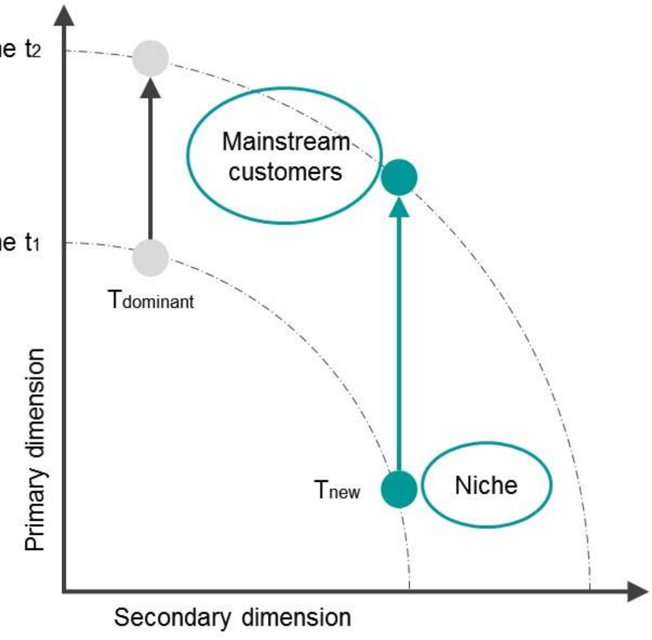

Figure 3: Theory of disruptive innovations (Sood & Tellis, 2011, p. 341)

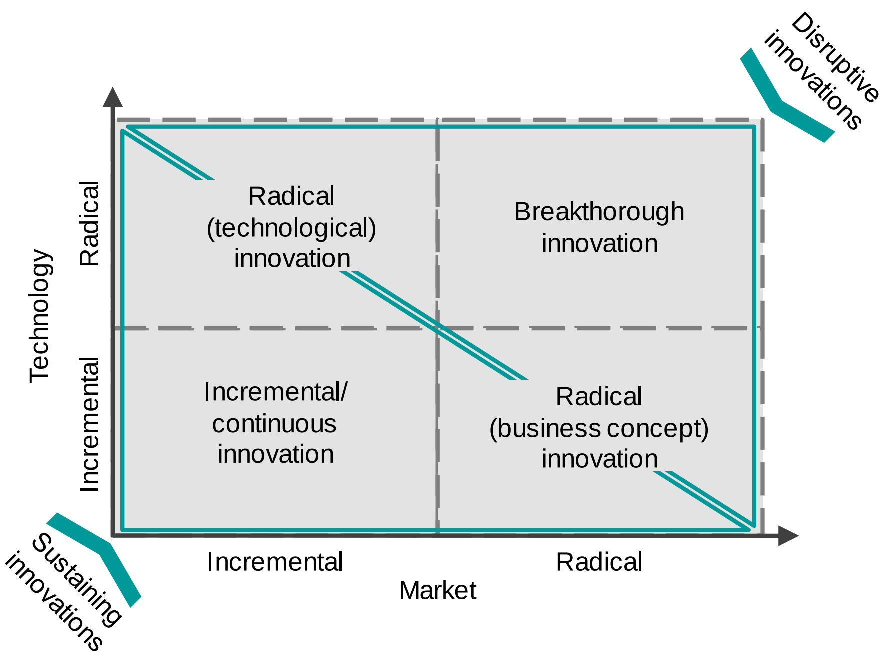

Figure 4: Classification of an innovations degree of newness (Assink, 2006, p. 217; Christensen, 1997/2011, p. xviii; Norman & Verganti, 2014, p. 89)

Potential sources of innovation

It is of fundamental importance for a corporation to know about potential sources of innovation, to actively take advantage of them and to achieve a high innovation performance (Tödtling, Lehner, & Kaufmann, 2009). When reflecting innovations and their contribution to the corporate competitive advantage the question that often arises is: What are the specific sources of innovation a corporation can utilize?

Scholarly research argues that internal (Baark et al., 2011; Chesbrough, Vanhaverbeke, & West, 2006) as well as external (Chesbrough et al., 2006; West, Salter, Vanhaverbeke, & Chesbrough, 2014) innovation sources contribute to the innovativeness of the firm and thus to the corporation's competitive advantage. Utilizing internal, as well as external sources for corporate innovations, is known under the term open innovation (Chesbrough et al., 2006). However, corporate innovations come predominantly from outside the firm (Laursen & Salter, 2004; Varis & Littunen, 2010; West et al., 2014; West & Bogers, 2014) and are therefore a product of changing corporate environmental conditions (Damanpour & Gopalakrishnan, 1998).

Hippel (1988) was the first who used the term sources of innovation. He defined sources of innovation as "*categorizing firms and individuals in terms of the functional relationship through which they derive benefits from a given product, process, or service innovation*" (Hippel, 1988, p. 3). On the contrary, Drucker (1998) described sources of innovation as the reason that leads to the opportunity of the innovation. Other authors referred to sources of innovation to the locality or geographic occurrence (Graf, 2016). However, the majority of academic scholars, such as West and Bogers (2014), Varis and Littunen (2010), (West et al., 2014) or (Laursen & Salter, 2004) define sources of innovation with regard to its origin, such as the institution or person originating the innovation. This can be considered a more thorough definition of Hippel's early description.

The present thesis considers startups and entrepreneurial ventures as a source for corporate innovation. Thus, this part of the review will focus on utilizing external sources for innovation. Table 4 shows the variety of external sources of innovation identified in the literature. Furthermore, West and Bogers (2014) identified several possibilities on how external sources of innovation can be leveraged. Firstly, firms can identify or search for innovation from external stakeholders or specialist with useful knowledge, such as suppliers (Schiele, 2010), customers (Grimpe & Sofka, 2009), competitors (Lim, Chesbrough, & Ruan, 2010), universities (Cassiman, Di Guardo, & Valentini, 2010) etc. Secondly, a company can utilize external search via technology scouts (Rohrbeck, 2010b), intermediaries (Jeppesen & Lakhani, 2010) or online communities (Füller, Matzler, & Hoppe, 2008), for example via crowdsourcing. Thirdly, the corporation can enable and filter information from external sources through an appropriate mechanism to incentivize, by organizing innovation contests (Boudreau, Lacetera, & Lakhani, 2011; Terwiesch & Xu, 2008), or by addressing more intrinsic motivational characteristics for example via open source structures (West & Gallagher, 2006). Lastly, external innovation can be licensed (Laursen, Leone, & Torrisi, 2010) or acquired (Dushnitsky & Shaver, 2009).

Type	Source of Innovation
Internal	Inside the company
External	
Market	Suppliers
	Clients, customers, users, individuals, brand community members
	Competitors
	Consultants
	Commercial laboratories/R&D enterprises
	Substitutors
	Complementors
	Startups
Institutional	Universities or other higher education institutes
	Government research organizations
	Other public sector e.g. business links, government offices
	Private research institutes
	Non-profit research institutions
Other	Professional conferences, meeting
	Trade associations
	Technical/trade press, computer databases
	Fairs, exhibitions
	Patent disclosure
	Press, media
Specialized	Technical standards
	Health and safety standards and regulations
	Environmental standards and regulations

Table 4: Sources of Innovation (Baark et al., 2011, p. 9; Bogers & West, 2012, p. 68; Cassiman et al., 2010, p. 884; Clausen, Pohjola, Sapprasert, & Verspagen, 2012, p. 18; Grimpe & Sofka, 2009, p. 495; Laursen & Salter, 2004, p. 1202; Lim et al., 2010, pp. 298–299; Schiele, 2010, pp. 138–139; Weiblen & Chesbrough, 2015, pp. 87–88)

Despite the extensive academic contribution on different sources of innovation as well as the widely-recognized importance for innovation, one important source has not been recognized in the literature of innovation sources yet: startups. Startups are considered as agile, innovative, creative, fast and likely to take larger risks and therefore more likely to develop disruptive innovations (Weiblen & Chesbrough, 2015). Corporations begin to integrate entrepreneurial firms as a valuable source for innovations (Weiblen & Chesbrough, 2015) but, as the preceding analysis of the literature showed, startups have not yet attained considerable attention in the academic literature on serving as such, despite their recognition in the corporate entrepreneurship and open innovation literature.

Startups are recognized as an opportunity to internalize external innovation sources and increase the corporate innovation performance. However, the identification of these external sources relies on scouts, coincidence or self-application or -presentation of potentially interesting startups (Pauwels et al., 2016). Currently, there exists no structured and automated approach to proactively identify opportunities within the startup context or to utilize startups as a source of environmental information and change.

Conclusion 9: The academic literature on potential sources of innovation provides an extensive overview of potential sources. However, startups and entrepreneurial firms are not explicitly considered as a potential source in the academic literature yet, despite their benefits and superiority in many aspects relevant for a company's innovativeness.

2.3.2 Startup

Definition of a startup

The academic literature provides only a few definitions about startups. Besides the term startup, there exist a few related or synonymously used phrases such as new venture, small firm, small business, entrepreneurial firm, entrepreneurial venture (Carland, Hoy, Boulton, & Carland, 1984; Oviatt & McDougall, 1994; Unger, Rauch, Frese, & Rosenbusch, 2011). Therefore, this chapter will provide an overview of different definitions and design a comprehensive definition of startups.

The origin of the definition of startup goes back to Schumpeter's definition of an entrepreneur (Carland et al., 1984). Carland et al. (1984) elaborated on the characteristics described by Schumpeter and defined an entrepreneurial venture as a:

> *"[…] venture that engages in at least one of Schumpeter's categories of behavior: that is, the principal goals of an entrepreneurial venture are profitability and growth and the business is characterized by innovative strategic practices"* (p.358).

Gilbert, McDougall, and Audretsch (2006) has defined the term new venture as a firm which is subject to the liability of newness and smallness. Luger and Koo (2005) are the first to explicitly define the term startup as:

“[…] a business entity which did not exist before during a given time period (new), which starts hiring at least one paid employee during the given time period (active), and which is neither a subsidiary nor a branch of an existing firm (independent)” (p.19).

Blank (2013) chooses an entirely different set of characteristics as important for a startup and refers to it as a “*[...] temporary organization designed to search for a repeatable and scalable business model*” (p. 5). Dee, Gill, Weinberg, and McTavish (2015) extend the definition of Blank (2013) by adding to its characteristics that it is growth oriented with regard to employees, revenue and customers.

Moreover, Berends et al. (2014) appends that startups are flexibility, characterized by direct decision making and able to communicate efficiently. Another important characteristic of a startup is its age. The definition in the literature range from three over five to ten years (Graham, 2013; Hill, 2015; Ripsas & Tröger, 2014). This research will consider ten years as the maximum age of a startup.

In this report, a startup is defined as a business entity that aims to design a repeatable and scalable business model and characterized by newness, flexibility, direct decision-making, and the ability to respond quickly to market changes. The maximum corporate age of a startup is ten years.

Startups as a potential source for corporate innovations

The literature about sources of innovation does not include startups as a potential source. Thus, the utilization of a startup as a source of innovations will be specifically examined in the subsequent chapter by reviewing literature on startups and entrepreneurship.

Startup firms differ significantly to incumbents. Startups and new ventures have several advantages over incumbent firms. They posess better abilities to identify opportunities (van Burg, Podoynitsyna, Beck, & Lommelen, 2012), show greater flexibility, have less beaurocratic structures, efficient internal communication patterns and establish strong relationships with their partners faster (Berends et al., 2014; Fiegenbaum & Karnani, 1991). Moreover, startups are able to respond to changing market conditions quickly (Crockett, McGee, & Payne, 2013; Rothwell, 1984). Furthermore, Plehn-Dujowich (2013) outlined several advantages of small and/or young businesses compared to large and/or old businesses. By analyzing the number of products per employee, Plehn-Dujowich (2013) showed that on average small firms launched significantly more products per employee than large firms. Companies with a median employment of 1000 persons, had 0.0214 products per employee, companies with a median staff of 375 showed 0.0534 product per employee and firms with a median employment of 15 showed a product employee ratio of 0.8767. This comparison of small and large firms shows that small firms launch significantly more products per employee than large firms. However, this comparison does not consider for example the level of complexity of the products.

Taking patent performance as another measure for innovativeness between small and large firms, Isom and Jarczyk (2009) outline that small firms are 13 to 14 times more productive in developing innovations, based on a patent to employee ratio. By looking more closely into the performance of patenting as well as research and development, Plehn-Dujowich (2013) showed that small as well as young firms have a better performance than large and old firms. The performance measure has been based on patent citations and the cumulative sum of R&D expenditure (Plehn-Dujowich, 2013). The study evaluated that small firms have 2.21 times higher patent citation per dollar of R&D stock than large firms. Moreover, young firms have a 2.27 times higher citation per dollar of R&D stock than established firms. By narrowing down the focus it became obvious that young, small firms have a 2.4 times better citation to R&D stock ratio than young, large firms. In the group of old firms on the other hand, small firms have a 2.0 times higher productivity than large ones. Finally, among small firms, young firms achieve 2.5 times better patent citation to R&D stock ratio (Plehn-Dujowich, 2013). In summary, small as well as young firms show a better innovation performance compared to established and/or large firms.

Additionally, Chandy and Tellis (2000) outlined that small entrepreneurial firms tend to introduce more radical innovations than incumbents, which rather tend to improve their existing product portfolio by innovating incrementally. Larger firms show the tendency to innovate less than small and new ventures, which often create, define, discover and exploit potential opportunities ahead of large firms (Zahra et al., 2006). Assink (2006) adds, that large companies are less creative than small firms and therefore develop less innovative ideas. This line of argumentation is also supported by empirical evidence of other research scholars. They found out that startups have more radical or disruptive innovations than incumbents (Henkel et al., 2015; Reid et al., 2015; Tidd, 2001). Moreover, startups are more likely to commercialize their radical innovations successfully (van Praag & Versloot, 2007). However, there is also contrary evidence on the development of disruptive innovations by small firms which indicate that incumbents and entrants are introducing disruptive innovations on an equal basis (Sood & Tellis, 2011).

Several scholars outline that important informantion come from outside the firm (Cohen & Levinthal, 1990; Horváth & Enkel, 2014). Weak signals are an indication of environmental change which in turn is an important factor to create new innovations (Yoon, 2012). These innovations often come from pioneers or special groups rather than domain experts (Heuschneider & Herstatt, 2016).

Therefore, startups are considered as a perfect source to look for environmenal information, regarding change in the sphere of product, service, technology, or business model.

Furthermore, startups are considered as a source of potential corporate innovations. This has also been concluded by other research scholars, such as Dushnitsky and Shaver (2009), Kortum and Lerner (2000) or Alvarez and Barney (2001). Moreover, it is emerically supported by the work of Block and Keller who showed that the importance of innovations coming from large corporations diminished throughout the past 35 years (Block & Keller, 2009).

By expanding the view outside the usual boundaries, e.g. outside the industry or outside the usual sources for innovation, like those not directly related to the incumbents core business a company can lower the risk of blind spots and thus increase the potential to detect and discover potential future developments (Day & Schoemaker, 2004a; Paliokaitė & Pačėsa, 2015).

Futhermore, by exposing the company to novel ideas, pioneering technologies, different organizational structures and problem solving approaches might enable the incumbent firm to create more radical and breakthrough innovations itself (Ahuja & Morris Lampert, 2001). Ahuja and Morris Lampert (2001) mention two reasons for this phenomena. First of all, it enables a corporation to stay heterogeneous in its problem solving approaches. Secondly, through studying new technologies, it brings fresh thoughts into existing cognitive structures and cause-effect relationships (Ahuja & Morris Lampert, 2001).

Based on the aforementioned reasons why startups have some advantages over incumbent firms and why there are benefits for incumbent firms to consider startups as a potential source for coroprate innovations, scholars, policy makers as well as practitioners have just recently recognized the importance to support these new ventures and utilize them to create innovations, e.g. in incubation or acceleration programs (Mian, Lamine, & Fayolle, 2016).

Taking a closer look at the incubation and acceleration activities, Mian et al. (2016) showed that from 1980 to 2012 there was a huge rise in business incubation programs in the United Stated, from 0 to around 1250. On a global scale there exist currently around 7000 incubation programs (Mian et al., 2016).

In Germany are currently around 350 active business incubators (BVIZ, 2016). By looking into the German automotive industry specifically, it can be seen that the developments of the recent years show the recognition of the issue by the automotive industry and their intensified effort to cooperate with startups. Table 5 provides an overview of different acceleration, incubation or venture capital programs that are currently established in the automotive industry to identify and support potential startups.

Company	Name	Program	Start Year	Source
Porsche AG	Porsche Digital	Venture capital	2016	Porsche AG. (2016). Porsche Digital GmbH. Retrieved from https://newsroom.porsche.com/de/unternehmen/porsche-kompetenz-zentrum-digital-gmbh-12569.html
Daimler AG	Startup-Autobahn	Accelerator	2016	Daimler AG. (2016). Startup Autobahn – No Limits. Retrieved from http://www.startup-autobahn.com/de/
Volkswagen AG	Digital Lab	Corporate venturing	2016	Volkswagen AG. (2016). Digital Lab. Retrieved from http://www.volkswagenag.com/content/vwcorp/info_center/de/news/2016/03/Silicon_Valley.html
BMW AG	BMW Startup Garage	Accelerator	2015	BMW AG. (2016). BMW Startup Garage. Retrieved from http://www.bmwstartupgarage.com/
Robert Bosch GmbH	Robert Bosch Start-up GmbH	Incubator	2014	Robert Bosch GmbH. (2016). Robert Bosch Start-up GmbH. Retrieved from http://www.bosch-startup.com/
BMW AG	BMW iVentures	Venture capital	2011	BMW AG. (2016). BMW i Ventures. Retrieved from http://www.bmw.com/com/en/insights/corporation/bmwi_ventures/index.html

Table 5: Accelerators, incubators and venture capital companies

Weiblen and Chesbrough (2015) outline the potential benefits a collaboration between startups and incumbents as a perfect prerequisite to enhance corporate innovativeness. Startups have ability to react to environmental change, whereas incumbents contribute their available resources (Weiblen & Chesbrough, 2015).

The collaboration or interaction between startups and corporations can thereby take different forms, such as corporate venturing, corporate incubation, startup program (outside-in), startup program (platform), as outlined by Weiblen and Chesbrough (2015).

Weiblen and Chesbrough (2015) differentiate between two basic perspectives of an engagement between corporations and startups, an outside-in and an inside-out perspective. The outside-in perspecive focuses on a facilitation of innovations through respective startups and an internalization of promising ideas by putting the startup into the role of a supplier. The Inside-out perspective follows a contrary approach and requires as well as faclilitates the startup to use corporation supplied technologies and expand the market of the parent organization.

Figure 5 provides a short overview of these four basic engagagement models between startups and corporations. Another and rather more traditional engagement possibility for large corporations to engage with startups is via an acquisition (Kohler, 2016; Weiblen & Chesbrough, 2015).

		Direction of innovation flow	
		Outside-in	Inside-out
Equity involvement	Yes	**Coporate venturing** Participate in the success of external innovation and gain strategic insights into non-core markets.	**Corporate Incubation** Provide a viable path to market for promising corporate non-core innovations
	No	**Startup program (Outside-in)** Insource external innovation to stimulate and generate corporate innovation	**Startup program (Inside-out)** Spur complementary external innovation to push an existing corporate innovation

Figure 5: Engagement models between corporations and startups (Weiblen & Chesbrough, 2015, p. 81)

The academic literature describes several possible mutual benefits that arise from a collaboration between incumbents and startups. The primary goal is of course the enhancement of the companies innovativeness by utilizing startups as a source of new ideas and an instrument for exploration (Pauwels et al., 2016), by harnessing their entrepreneurial power (Kohler, 2016). More specifially, the engagement with startups contributes to the corporation by bringing new knowledge (Dushnitsky & Shaver, 2009), new technologies, practices or market insights (Siegel, Siegel, & MacMillan, 1988), as well as new approaches to solve problems (Ahuja & Morris Lampert, 2001) into the incumbent firm. Startups on the other hand receive help for the execution by gaining access to necessary resources, such as human, financial and physical ones as well as knowledge over the existing product (Alvarez & Barney, 2001; Kohler, 2016). Moreover, they have not such a high reputation, reliability and legitimacy due to the few years of experience (Baum, Calabrese, & Silverman, 2000).

Nevertheless, the benefits may depend on the specific context the corporation and the startup are in. Stettner and Lavie (2014) for example investigated that it is more beneficial to use acquisitions to attain distinct knowledge and explore for new technologies, product or business models while the internal knowledge is better utilized to exploit and thus refine and improve existing products and services. Basu and Wadhwa (2013) further investigated acquisitions and empirically proved that they are more likely to lead to radical or disruptive innovations than alliances or corporate venture capital activities.

As described above, the academic literature outlines quite extensively that incumbents should engage with startups to increase corporate innovativeness and thus gain a competitive advantage. Furthermore, it outlines several possibilities on how to engage with startups and more specifically some work even investigates the benefits of different engagement models under different environmental conditions.

However, what has not been addressed in the academic literature is how to identify startups that are potentially valuable for the corporate firm. There are different proposed methodologies on how weak signals (Yoon, 2012), environmental information (Seo et al., 2016), white spots (Yoon et al., 2013), technological opportunities (Yoon et al., 2015) or ideas (Thorleuchter & van den Poel, 2015) can be identified (see Chapter 2.2.8), nevertheless there are no contributions on how to identify startups, neither is there a thorough proposition that startups could serve as a source of environmental information. The identification does thereby refer to the methodological approach to recognize or spot startups that are of potential value for the incumbent firm. This however is not yet addressed in academia, neither in the corporate foresight literature, nor in the innovation management, open innovation or corporate venturing literature.

Two possibilities to identify startups have been mentioned by Kohler (2016) as well as Pauwels et al. (2016). They state that acceleration programs are characterized by an open application process. Although this open application approach has certain advantages, it limits the choice of potential startups to those actively applying and does not scan the market for those that might be most beneficial for the corporation or yield the highest market potential. Some acceleration programs go one step further and use dedicated events to scout the respecting startups (Pauwels et al., 2016), nevertheless also here, startups have to apply or be identified and selected for such an event. This drawback has been recognized by researches in the field. Hathaway (2016), for example, proposes to use investor or startup databases to identify accelerator programs in his research. Moreover, Remane et al. (2016) utilized a startup database to identify startups in order to assess new digital business models types in the mobility sector. Thus, an investor or startup database could also serve as a source for startups in general.

Conclusion 10: Based on the outlined benefits of startups and an engagement between startups and established corporations it should be a prerequisite for corporations to incorporate startups into their corporate foresight activities. Incumbents who are in search of innovations, specifically those of a radical or disruptive nature, should actively engage with startups, utilize their advantages and internalize external innovation. This is supported in the recent developments in the field of corporate venturing and open innovation as startups are already considered a highly relevant source for corporate innovations there.

2.4 The connected vehicle

Current developments indicate that the future car will be connected to its environment to a much greater extent. Connectivity is not only a market demand, but it is also a prerequisite as well as an enabler for the automated and the electric driving (Slama, Puhlmann, Morrish, & Bhatnagar, 2015). In the light of this, the connected vehicle has the potential to transform the whole mobility sphere (Minelli, Izadpanah, & Razavi, 2015).

The connected car originates from the wider concept of the internet of things (Johanning & Mildner, 2015; Rayes & Salam, 2017), which in turn has its origins in the concept of machine-to-machine communication (Slama et al., 2015). The most basic definition of the internet of things refers to it as networkable things that are either a physical object or a virtual entity (Lee, Crespi, Choi, & Boussard, 2013). Based on the definition of the internet of things, we can characterize the connected car as a car that communicates and networks with other physical objects or virtual entities.

Gordon (2016) describes the connected vehicle on a very basic level as a car that connects to an external network. Swan (2015) gives a more sophisticated definition. She describes the connected vehicle as a car that is in possession of internet access and allows the car to communicate with its environment. Hong, Shin, and Lee (2016) emphasizes a different point about the connected car, according to their definition it is a vehicle that allows the driver as well as the passengers to access, consume and share information through the communication system of the vehicle, e.g. via vehicle-to-vehicle or vehicle-to-infrastructure communication (Gora & Rüb, 2016; Hong et al., 2016).

Lu, Cheng, Zhang, Shen, and Mark (2014) provide the most sophisticated definition and defined the connected car as the following:

> "*Connected vehicles refer to wireless connectivity-enabled vehicles that can communicate with [their] internal and external environments, i.e., supporting the interactions of vehicle-to-sensor on-board (V2S), vehicle-to-vehicle (V2V), vehicle-to-road infrastructure (V2R), and vehicle-to-internet (V2I). These interactions, establishing multiple levels of data pipeline to in-vehicle information systems, enhance the situational awareness of vehicles and provide motorist/ passengers with an information-rich travel environment*" (p.289).

Based on the preceding definitions it can be said that connected vehicles collect, process, interpret, share and utilize all incoming and outgoing data, to actively communicate with vehicles, infrastructure, and other technological devices or things (Koslowski, 2012; Lu et al., 2014).

The communication could thereby be about the car’s position, speed, routes, lane changes, or stopping intentions (Gora & Rüb, 2016). These cars could also agree to common driving strategies and thus lead to a smoother overall traffic flow (Gora & Rüb, 2016). In order to achieve the connection of the car to its outside environment, it requires new hardware and software inside as well as outside the vehicle (Johanning & Mildner, 2015).

In general, the connected vehicle consists of 3 main domains, the vehicle with its in-vehicle network, electronic control units (ECUs) and onboard diagnostic tools (OBD) (Johanning & Mildner, 2015; Kleberger, Olovsson, & Jonsson, 2011), the vehicle’s communication unit, called communication control unit and lastly services established around the connected vehicle (Kleberger, 2015). The ECUs control the in-vehicle network and information flow, receiving information from sensors process them and pass them on to the actuators to perform the required tasks (Lu et al., 2014). The communication control unit connects the vehicle to other devices and manages the communication between the car and the other devices (Kleberger, 2015). The services that are already establishing around the connected car and those that are expected to develop are manifold. The consulting firm Everis (2015) distinguished between 5 different types of services, traffic safety, connected infotainment, traffic efficiency, cost efficiency, as well as convenience and interaction. Everis (2015), Gordon (2016) and Slama et al. (2015) proposed several potential services which will be outlined in the following. Firstly, services related to traffic safety are for example ecall, an emergency service, roadside assistance or breakdown services, stolen vehicle assistance or remote diagnostics. Secondly, connected infotainment includes multimedia streaming, social media or car WiFi networks. Thirdly, traffic efficiency refers to services such as traffic information, navigation, fuel prices or parking. The fourth group of services relates to cost efficiency for example usage-based insurance, driver behavior monitoring, electric vehicle charging, remote maintenance or the integration of the electric vehicle into the smart grid. Fifthly, and lastly, are services related to convenience and interaction with examples like remote control of the vehicle, such as heating or engine start, car sharing, intermodal services or electronic toll.

The whole transformation of the current vehicle to a more connected one, with all the required technological, product, service and business innovation that are already taking place and expected to develop in the future are an enormous potential for current car manufacturers to gain access to new business fields. However, it also poses the risk of being outpaced by competitors and new market entrants in critical areas of the connected vehicle.

Due to that reason, it can be concluded that the automotive industry is expecting major changes during the upcoming years and that it is vital for every manufacturer to sense the indicators of that change. Furthermore, it is necessary to look for interesting ideas and potential innovation and actively pursue change internally as well as externally to not lose ground in this competitive environment.

3 Methodology

The methodology of this thesis is of an exploratory nature, also known as grounded theory methodology. The research utilizes precedent research, carefully interprets these and transforms them into a new exploratory research approach, to address the actual underlying research question (Rowe, 2014). As proposed by Cooper and Schindler (2014) a study of an exploratory nature is particularly suitable for investigative research topics that have not been addressed in the academic literature previously. Based on primary as well as secondary data, the exploratory research study proposes and tests a model to identify startups that are of potential relevance to the corporation and provide valuable input information for the corporate foresight of the firm.

This is considered a valuable contribution as it combines the method that has been applied in other contexts (see Chapter 2.2.8), to identify weak signals (Yoon, 2012), environmental information (Seo et al., 2016), white spots (Yoon et al., 2013), technological opportunities (Yoon et al., 2015), or ideas (Thorleuchter & van den Poel, 2015) with the identification of startups, addressing the suggestion and requests from the corporate foresight and corporate venturing literature.

Firstly, the methodology enables corporations to utilize startups within the corporate foresight process in a more structured and sophisticated way as a source of environmental information and/or innovation. Secondly, it addresses the request that business opportunities and knowledge need to be actively identified (Maine, Soh, & Dos Santos, 2015; Spithoven et al., 2010) and that the corporate foresight literature needs to develop more sophisticated search strategies to identify discontinuities, trends and innovations (Heuschneider & Herstatt, 2016). Thirdly, it incorporates the suggestion of Hathaway (2016) as well as Remane et al. (2016), who proposed to utilize a startup or investor database to identify accelerators or startups as a way to improve the identification procedure.

3.1 Research Design

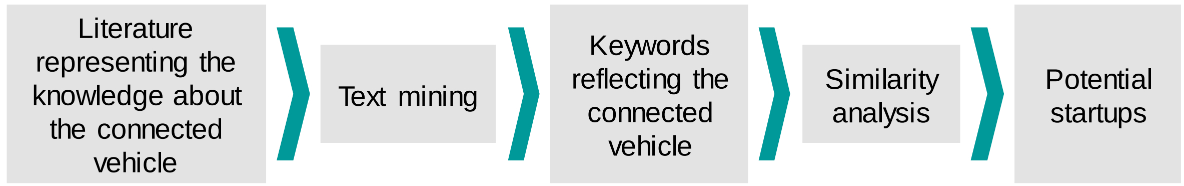

Figure 6: Research model (Own Illustration)

The research design was developed, addressing the first four steps of the general corporate foresight process, namely: scoping, gathering information, scanning and interpretation, (see Figure 1). As described earlier in Chapter 2.2, the discovery of innovative ideas as well as corporate environmental change is vital for a company. Startups are able to address both, as they are particularly innovative, full of ideas and respond with their innovations to environmental change, filling the gap that has been opened and left behind by unfulfilled market demands, technological or business model opportunities (van Burg et al., 2012; Weiblen & Chesbrough, 2015).

The research model as depicted in Figure 6 aims to propose a methodological approach to identify startups that are of potential relevance for a corporation based on a selected field or topic. Relevance does thereby address the potential value for the corporation as the startup might yield important environmental information or valuable innovations.

Overall the aim of the methodology was to employ the proposition of Hathaway (2016) as well as Remane et al. (2016) and utilize a startup database to identify startups. In order to identify these startups from the database it is necessary to have a certain set of keywords that represent the topic of interest and which can be used to search in the database for matches.

In general, the extraction of keywords, referred to as single words, terms, phrases or segments which appropriately represent the content of a document (Onan, Korukoğlu, & Bulut, 2016; Siddiqi, 2015), can be performed either manually or automatically. As with other foresight methods (Daheim & Uerz, 2008; Yoon, 2012; Yoon et al., 2015), a manual extraction would require a large amount of human resources with expert knowledge in the respecting field. In order to avoid this potential constraint, an automatic keyword extraction via a text mining approach is proposed. The value of an automated keyword extraction via text mining lies primarily in the possibility to process and analyze large amounts of unstructured text data and identify the main content, represented by keywords (Seo et al., 2016). As indicated, text mining has the potential to mitigate the drawback of an expert based approach as it has the benefit to reduce the number of experts and simultaneously augment and amplify the capabilities of the remaining experts, while minimizing the costs, compared to a manual procedure (Siddiqi, 2015;

Zhang et al., 2016). Additionally, it also enables to overcome expert biases, empower the detection of new facts, patterns, and knowledge that are not perceived by experts due to a limitation of their knowledge base or a prejudice of their vision and believes (Miranda Santo, Coelho, dos Santos, & Filho, 2006). After the extraction of the keywords, they were utilized to search the database for startups. This was implemented by looking for occurrences of the keywords in the startup descriptions and thus measuring the similarity between the both.

Recent scholarly contribution has shown that text mining techniques are a valuable tool to extract relevant keywords from different sources (Thorleuchter & van den Poel, 2015; Yoon et al., 2015; Zhang et al., 2016). Yoon (2012) used text mining techniques to derive weak signals from web news articles, based on the frequency of the respecting keywords. Kostoff et al. (2004) applied a literature based text mining for the identification of disruptive innovations. More specifically, they used text mining to detect technology components and further, utilized experts for identifying potentially disruptive innovations. Thorleuchter and van den Poel (2015) proposed an approach to use text mining from documents or websites to mine the text for keywords with the ambition to identify ideas that might serve to solve a given problem and subsequently derive weak signals from these ideas. Furthermore, Zhang et al. (2016) utilized text mining and keyword extraction techniques to identify hot topics in science, technology and innovation over a five-year period and used the acquired information to create a sophisticated technology roadmap.

A variety of possibilities exists for the performance of extracting respective keywords. These range from programming language based implementations and usage of open source algorithms, such as RAKE (Rose, Engel, Cramer, & Cowley, 2010) and Maui (Elhenfnawy et al., 2016), over web-based tools that offer basic analysis and visualization options such as Voyant Tools (Sinclair & Rockwell, 2016), to data analysis software such as RapidMiner (Kotu, 2014). RapidMiner provides ample text mining functionalities in a graphical user interface. In this research RapidMiner was employed for the processing of the data as it allows to use a large variety of text analysis, pre-processing and statistical tools while being easy to use as well as open source. The advantage of a graphical interface and a relatively simple application of the tool is not only an advantage for the execution of this research, but also for the use of the entire process in practice. It enables every corporate foresight or even a domain expert to perform the analysis as no specific programming capabilities are required.

It is an inherent characteristic of text mining that the quality of the output strongly depends on the quality of the input (D'Haen, van den Poel, & Thorleuchter, 2013). Because the text mining shall function as a supplement or even a replacement of expert resources, it is a prerequisite to perform the text mining based on expert literature. This is also considered necessary to ensure a certain quality of the information. Therefore, the respecting unstructured text documents used in this research were extracted from various professional sources, based on predefined search terms about the connected vehicle (see Chapter 3.2 for a more detailed description of how the data was collected).

Regarding specific text mining and extraction methods, there exist several methods on how to derive keywords from unstructured documents (Siddiqi, 2015). Siddiqi (2015) classifies existing text processing and keyword extraction methods (see Table 6).

Category	Description
Based on term weight	
Term count	Measures the number of times a term occcurs in a document.
Term frequency	Number of times a term occurs in a corpus.
Document frequency	If a word (w) occurs in a number of documents (n) in a corpus of (N) documents where (n) < (N), then (n) is called the document frequency of word (w).
Inverse document frequency	The more rare a word the higher its importance, IDF = log (N/dm). Where N = no. of documents in the corpus, dm = no. of documents containing the word m. A word which occurs in all documents of the corpus has the value 0 and has a low relevance. A word which occurs in only one document has the value 1 and is of high relevance.
Average frequency	Ratio of the number of words in a document devided by the number of documents in a corpus.
Relative frequency	Ratio of the number of words in a document devided by the number of documents in a corpus that word occurs in.
Term length	Number of tokens in a keyphrase.
Term frequency - inverse document frequency	Is used to weigh the importance of a word in a corpus. It combines term frequency and inverse document frequency and thus leverages their drawbacks.

Table 6: Keyword extraction techniques (Kotu, 2014, p. 278; Siddiqi, 2015, p. 21)

For the extraction of keywords, the term frequency - inverse document frequency (TF-IDF) approach was utilized as it incorporates the advantages of both, the term frequency as well as the inverse document frequency methods (Yoo & Yang, 2015). The TF-IDF extraction technique is the most widely applied technique for text mining and keyword extraction and is commonly utilized and recognized in academic research (Peng, Liu, & Zuo, 2014).

The term frequency (TF) counts how often a term appears in a document. The $TF_{(p,d)}$ of a keyword phrase (p) in a document (d) can be expressed as the following (Siddiqi, 2015):

$$TF_{(p,d)} = \frac{freq(p,d)}{size(d)}$$

$TF_{(p,d)}$ is calculated by dividing the freq (p,d), which represents the occurrence frequency of phrase (p) in document (d), by the size (d), which expresses the number of phrases in document (d) (Siddiqi, 2015). The inverse document frequency (IDF) on the other hand is depicted as following:

$$IDF_p = log_2(\frac{N}{DF_p})$$

The variable (N) represents the number of phrases in the corpus and (DF_p) or document frequency the number of documents that include the term (Siddiqi, 2015). Finally, the TF-IDF can be expressed as:

$$TF - IDF = TF_{(p,d)} * IDF_p$$

Besides the TF-IDF calculation, an additional step is necessary to extract the keywords from the documents: clustering. Clustering is referred to the process of the identification of natural groupings within data or a set of documents (Kotu, 2014). Clustering is considered as an essential step as it enables to automatically divide the documents into clusters based on their content and subsequently extract the most relevant keywords per cluster, like other text mining research has shown in different areas (Thorleuchter & van den Poel, 2015; Yoon et al., 2015; Zhang et al., 2016).

The academic literature provides ample cluster algorithms, which were not considered in their breadth in this research. In general, a distinction between three types, density-based, hierarchical and partitional clustering and can be made (Tran, Wehrens, & Buydens, 2006).

Density-based clustering is utilizing local clustering criteria to identify areas of density and thus define these dense areas as clusters (Tran et al., 2006). Partitional clustering, also known as iterative relocation clustering on the other hand, is arranging data points between clusters based on an iterative relocation approach (Shraddha & Emmanuel, 2014). A third group of clustering algorithms is referred to under the term hierarchical clustering and rests on an agglomerative (bottom-up) or divisive (top-down) division of data sets into nested partitions (Shraddha & Emmanuel, 2014). As the examination of the best algorithm for clustering and keyword extraction is not the focus of the research, only the most widely applied ones K-means, K-medoids, and DBSCAN will be described shortly in more detail (Satapathy, Bhateja, & Joshi, 2017; Tran et al., 2006).

The K-means clustering algorithm is a partitional algorithm and one of the most widely used (Shraddha & Emmanuel, 2014). It functions by searching for prototype data points, and all remaining data points are assigned to these prototype data points (Kotu, 2014). These prototype data points represent the center of the cluster and are also called centroids. Within K-means this centroid represents the mean of the data points (Kotu, 2014). K-medoids on the other hand, which is also a partitional algorithm, bases its centroid calculation not on the mean data points but on those that represent the whole cluster best (Kotu, 2014). In both cases, it does not necessarily have to be a real data point. Following Figure 7 shows an example of a Voronoi Diagramm, each color as a visualization of the cluster and the black dot as its centroid, representing a K-means clustering approach (Kotu, 2014). Figure 7 also visualizes the primary reason of why the clustering of the documents and the derivation of the respecting keywords is a useful approach. This is valuable as each group of documents represents a certain topic (each color) which could hold a particular set of unique keywords that otherwise would be missed out.

The third algorithm, DBSCAN is part of the density based algorithms (Sreenivasulu, Raju, & Rao, 2017). It builds on the general idea that regions with a high density of data points reflect the existence of a cluster, while areas with low density indicate outliers or noise (Shraddha & Emmanuel, 2014).

Within this research, the K-means clustering algorithm was applied as it is the most commonly used algorithm. This work does not claim to use the most comprehensive method, but rather aims to outline an approach on how to automatically identify keywords about a topic that can be utilized in a later step to determine startups in the respecting area of interest.

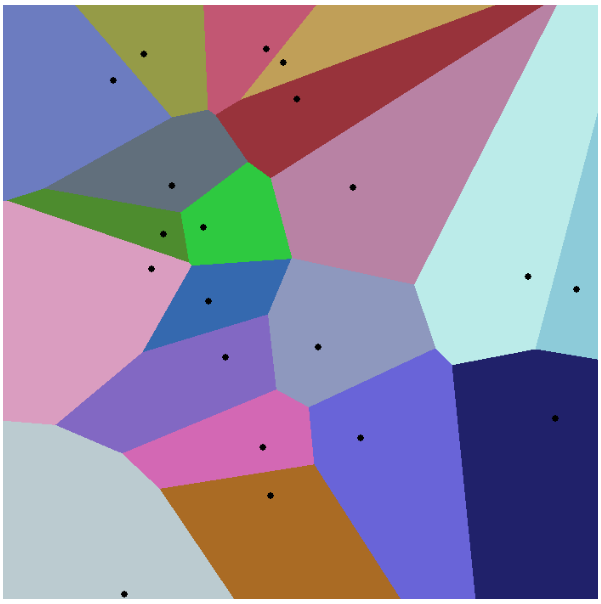

Figure 7: Euclidean Voronoi Diagram (Raincomplex, 2013)

Despite its frequent application, the K-means as well as other clustering approaches show some drawbacks. For K-means as well as K-medoids it is necessary to manually define the number of clusters (RapidMiner, 2016). To examine the most beneficial number of clusters the academic literature suggests to have the smallest possible value for intra-cluster distance and the largest possible value for inter-cluster distance (Yoon & Park, 2004). The average intra-cluster distance can be calculated by measuring the average distance between the centroid and all documents assigned to the respecting cluster (RapidMiner, 2016). The inter-cluster distance is part of the Davies-Bouldin index (DBI) and will be used as a second measure. The DBI calculated as the ratio between the intra-cluster distance and the inter-cluster distance (Kotu, 2014; RapidMiner, 2016). This means that the higher the inter-cluster distance and the lower the intra-cluster distance, the lower the overall DBI. The lower the DBI the better is the clustering (Kotu, 2014).

To further evaluate if the number of clusters (k) has an impact on the extracted keywords at the end of the text mining and keyword extraction, it is proposed to compare the top 20 keywords with a number of clusters of k=10 and the top 10 keywords with a number of clusters k=20. If there is an increased quality with a higher or lower number of clusters there should also be a difference in the keywords. After the proposed clustering was examined the keywords were extracted.

Finally, the extracted keywords and the startups from the startup database were checked for similarity, in order to identify those startups that fit the keywords best. These were expected to be startups in the field of the connected car or a related topic. Under the theoretical considerations outlined in Chapter 2.3, it is assumed that these startups have a potential to be of a disruptive nature and are valuable for a corporation with regard to environmental information as well as potential innovations. Again, the academic literature provides several possibilities to calculate the similarity between documents or data points, such as correlation similarity, simple matching coefficient, Jaccard similarity, or Cosine similarity (Kotu, 2014). To keep this step of the analysis simple, a procedure similar to the simple matching coefficient was proposed to measure the similarity between keywords and startup descriptions. An example of a startup and the associated description can be seen in Table 7. The startup description is a short description of the companies main value proposition or activities.

Startup name	Startup description
otonomo	otonomo provides a cloud-based software as a service (SaaS) solution for car manufactures to share and monetize car data and offer drivers in-car services.

Table 7: Example of a startup company and the associated description (http://www.crunchbase.com) (Crunchbase, 2016)

The simple matching coefficient usually measures the simultaneous occurrence of binary data points (Kotu, 2014). In this research, it was proposed to use a simple matching analysis between keywords and the short company description by examining the occurrence of each keyword in each company description and calculating a similarity index based on the aggregated number of keyword occurrences. The startup description was attained from a business information database called Crunchbase (2016).

Lastly, after potential startups were identified an open coding via *'in vivo'* coding combined with axial coding was applied as suggested for grounded theory methodology (Saunders, Lewis, & Thornhill, 2011). The purpose was to structure the identified startups and identify different areas of startup activity. The results will be presented in a startup map, similar to a patent map, which is considered to enable the visualization of complex information (Lee, Kang, & Shin, 2015). The mapping of patents to visualize, forecast and plan patent is widely recognized in academia and was recently utilized by Lee et al. (2015) for the mapping of technology opportunities identified from patents or by Yoon et al. to visualize recommended products (2016).

3.2 Data collection

As outlined above, there are two data sets that have to be collected to perform the research as described. Firstly, the necessary documents for the performance of the text mining and the extraction of the keywords have to be collected. Secondly, the information about the startups is an essential data set, as it is necessary for the similarity analysis between keywords and startups as well as for the derivation of potentially relevant startups.

In order to find relevant articles that address the topic of the connected vehicle, the SpringerProfessional (2016) database was utilized to identify relevant documents. The database was searched for predefined relevant search terms either synonymous or related to the connected vehicle. These predefined search terms were chosen by looking at a holistic picture of the connected vehicle. This holistic picture of the connected vehicle was created by applying generic search strings that are synonymous or strongly related to the term connected vehicle, such as connected car, connected mobility, connected driving, vehicle to vehicle, or vehicle to infrastructure in order to look for relevant literature. For the search of the keywords the temporal limitation was set to the period of October 2014 to October 2016. The language of the documents was set to English, as otherwise, the keywords would have to be translated into English again, posing the risk of mistranslation. Furthermore, having both English and German documents would have distorted the following clustering via RapidMiner. As source type, professional journals, and book chapters were selected. Professional journals are written by professionals of the respecting field and focus on applied science (University of Twente Library, 2013). It is expected that they provide more sophisticated information than web news, as proposed by Yoon (2012) while not being focused on scientific methods and technical details. For the subsequent text mining and keyword extraction, the documents were extracted as pdf files. The dataset about the startups emanates from a business information database, named Crunchbase. It provides information about public and private companies and stores around 116,000 data sets (Crunchbase, 2016). The database is a self-reported database with more than 295.000 contributors (Crunchbase, 2016). This research got granted extended access to the database and thus has access to the information listed in Table 8.

Information from Crunchbase has been utilized previously for research purposes, e.g. to identify areas of startup activity in the social media domain (Ghezzi, Gastaldi, Lettieri, Martini, & Corso, 2016), to predict investors funding behavior (Liang & Yuan, 2016), to analyze the impact of patents and trademarks on VC funding practices (Zhou, Sandner, Martinelli, & Block, 2016), or to locate geographical hotspots of innovation (Graf, 2016).

Data	Description
Company name	Name of the organization
Domain	Company URL
Country code	Country the company is based in
State code	State the company is based in
Region	Region the company is based in
City	City the company is based in
Status	Operating, closed, acquired, initial public offering (IPO)
Short description	Gives a short description of the company
Category groups	Category the company is operating in
Employees	Number of employees
Funding rounds	Number of funding rounds
Funding amount	Aggregated funding amount
Founding date	Date the company was founded
First Funding	Date of the first funding
Last Funding	Date of the last funding
Closing Date	Date the company was shut down
Email	Email adress
Phone	Phone number
CrunchBase URL	Company URL on CrunchBase
Twitter URL	Company URL on Twitter
Facebook URL	Company URL on Facebook

Table 8: Crunchbase data categories (http://www.crunchbase.com) (Crunchbase, 2016)

3.3 Research process

The research process was based on the research design, as depicted in Figure 6 and was separated in three research sub-processes: [A] the identification of the literature corpus, [B] the text mining and keyword extraction and [C] the similarity analysis. The whole research process is depicted in Figure 9.

The research process in its whole was developed specifically and uniquely for this research project, addressing the identified gap, described in Chapter 1.1. However, it is a combination of methodological approaches previously applied in the academic literature in the field of text mining, opportunity discovery or foresight.

The identification of news articles takes the work of Yoon (2012) as well as Iglesias, Tiemblo, Ledezma, and Sanchis (2016) as a reference. They employed text mining based on web news articles to extract keywords. Iglesias et al. (2016) used the New York Times as a source, while Yoon (2012) used the ProQuest database to collect web news articles. Within this research, the utilization of articles from professional journals from the SpringerProfessional database was proposed. It was intended to rely on professional journal articles rather than web news as they are more likely to represent expert knowledge as they are written by experts in the field (University of Twente Library, 2013).

The text mining step was based on the work of Thorleuchter et al. (2010), Yoon et al. (2014), (Iglesias et al., 2016) and Zhang et al. (2016). All four publications employed the TF-IDF for data preprocessing. Iglesias et al. (2016) provides the most detailed methodological description by outlining that the term extraction will consist of the process steps outlined in Figure 8.

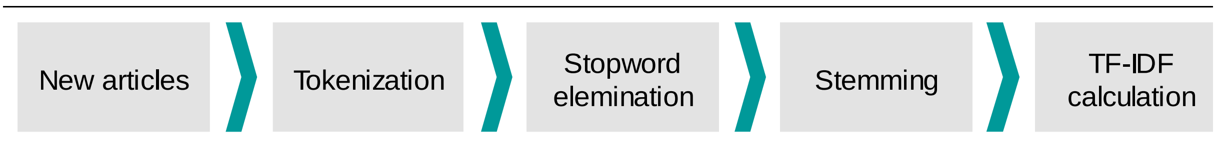

Figure 8: Structure of the term extraction (Iglesias et al., 2016, p. 93)

Furthermore, Zhang et al. (2016) applied a K-means clustering approach to group semantically similar documents, based on the TF-IDF values. This work applied the K-means based clustering approach but by using a more basic method to calculate the K-optimum, described in detail in Chapter 3.2. The similarity calculation is suggested as a general methodology for example by Kotu (2014) but has not been employed so far in the identification of environmental information of innovation.

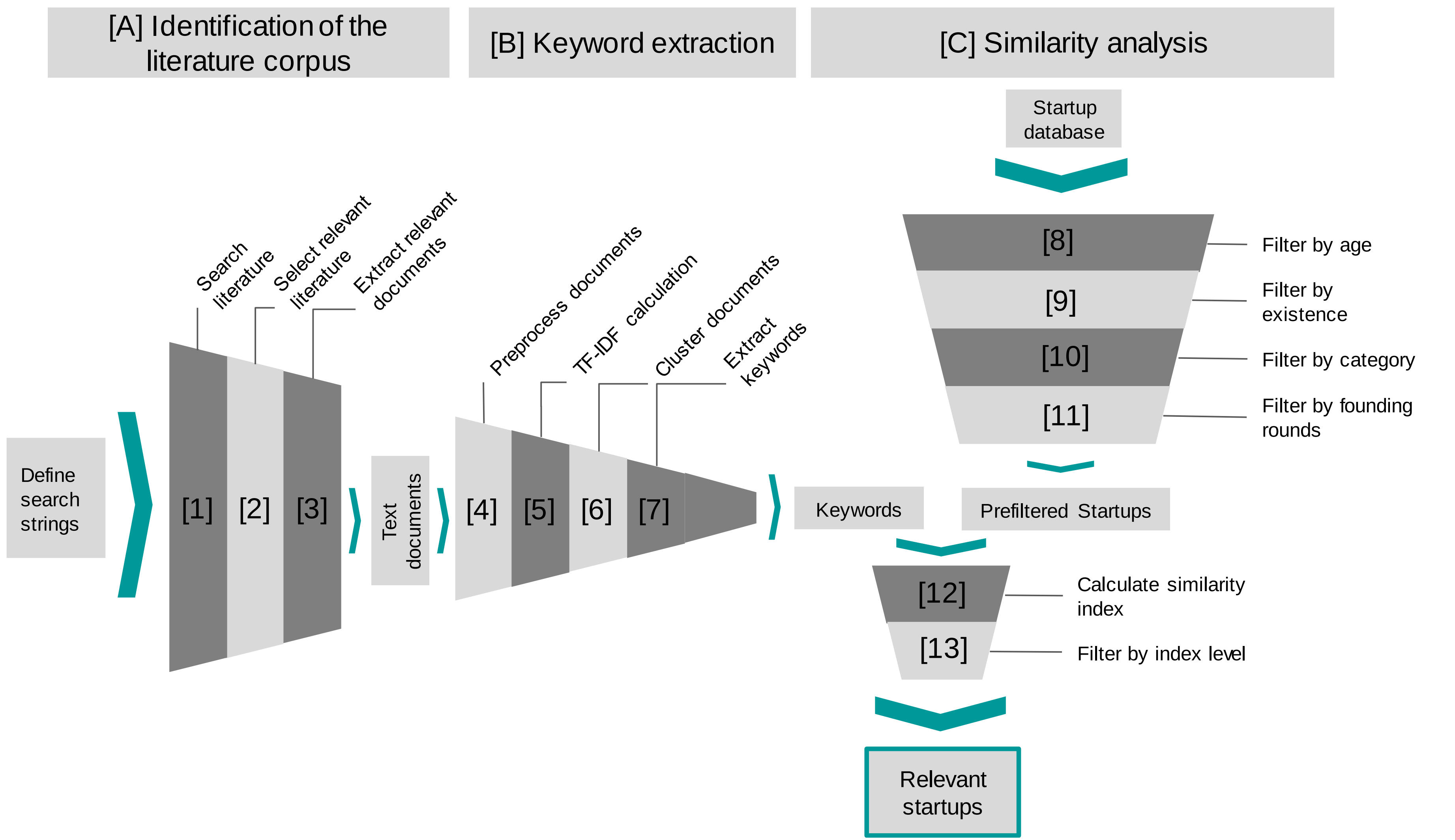

Figure 9: Visualization of the research process (Own illustration)

3.3.1 Identification of the literature corpus

The literature was derived as described in Chapter 3.2. Table 9 shows the search strings that were applied for books and professional as well as scientific journals. Furthermore, it lists the number of documents that were identified for each journal or search string as well as the number of documents that were considered as relevant (process part [A]).

Search string	Journal	Identified articles	Relevant articles
Journals			
connected vehicle	Auto Tech Review	119	15
	ATZ Worldwide	75	10
	ATZ elektronik worldwide	66	24
	International Journal of Automotive Technology	75	6
	Wireless Personal Communication	93	7
connected car	Auto Tech Review	89	20
	ATZ Worldwide	57	9
	ATZ elektronik worldwide	49	24
	International Journal of Automotive Technology	44	3
	Wireless Personal Communication	47	10
connected mobility	Auto Tech Review	53	11
	ATZ Worldwide	19	5
	ATZ elektronik worldwide	27	7
	International Journal of Automotive Technology	14	4
	Wireless Personal Communication	244	0
connected driving	Auto Tech Review	89	20
	ATZ Worldwide	53	3
	ATZ elektronik worldwide	52	6
	International Journal of Automotive Technology	55	1
	Wireless Personal Communication	43	5
intelligent vehicles	Auto Tech Review	61	10
	ATZ Worldwide	64	0
	ATZ elektronik worldwide	37	7
	International Journal of Automotive Technology	59	4
	Wireless Personal Communication	77	18
vehicle to vehicle	Auto Tech Review	17	5
	ATZ Worldwide	251	7
	ATZ elektronik worldwide	3	3
	International Journal of Automotive Technology	6	3
	Wireless Personal Communication	146	10
V2V	Auto Tech Review	10	5
	ATZ Worldwide	3	3
	ATZ elektronik worldwide	1	1
	International Journal of Automotive Technology	6	3
	Wireless Personal Communication	43	15

continued

Vehicle to infrastructure	Auto Tech Review	7	7
	ATZ Worldwide	5	3
	ATZ elektronik worldwide	1	1
	International Journal of Automotive Technology	4	2
	Wireless Personal Communication	26	5
V2I	Auto Tech Review	5	5
	ATZ Worldwide	2	2
	ATZ elektronik worldwide	1	1
	International Journal of Automotive Technology	3	2
	Wireless Personal Communication	31	6
Books			
connected vehicle		70	40
connected car		61	27
connected driving		5	1
connected mobility		7	1
	Total	**2375**	**387**

Table 9: Identified articles, grouped by search string and journal

In total, five professional journals and magazines were selected for the keyword extraction (section [1]). The selected journals are addressing a broad topic range while yielding a high number of publications over the selected time period. They were selected out of 101 journals as they yield the most search results, while not being focused on one particular topic or technical detail which might result in keyword distortions. Examples of such journals are the Journal of Mechanical Science and Technology, Mobile Networks and Applications or The International Journal of Advanced Manufacturing Technologies. Every article meeting the search results was opened and selected for further processing based on its relevance (section [2]). The pertinence was judged on its headline and abstract. Finally, in preparation for the text mining, all duplicates were removed, leading to a total amount of 387 documents. After eliminating all the duplicates that were found, 238 records remained in the corpus (section [3]).

3.3.2 Text mining and keyword extraction

In the second step of the research process (process part [B]), the previously identified documents that address various topics about the connected vehicle were text mined for the purpose of extracting relevant keywords that reflect the content of the entire literature corpus. The process of text mining and keyword extraction consists of four steps: [4] Preprocessing of the documents, [5] TF-IDF calculation, [6] clustering of the documents and the respecting keywords, [7] derivation of the top keywords per cluster (see Figure 9).

Section [4], the preprocessing of the documents, consists in a total of seven steps (see Figure 10). This is necessary to enable the following TF-IDF calculation and improve the quality of the outcome. Thorleuchter et al. (2010) followed a similar approach for preparing the input data in their research. Firstly, all documents were tokenized. The tokenization of the documents splits the text into a sequence of tokens, based on the defined splitting point, which can be based on non-letter characters or linguistic tokens (RapidMiner, 2016). In order to increase the quality of the tokenization, this research followed the proposition of Ertek, Tapucu, and Arın (2013) to utilize both tokenization operators. Subsequently, several stopword filters were applied. Filtering for stopwords is considered a necessary step as it filters out unnecessary, unwanted or meaningless words, such as prepositions, articles or manually defined words (Berry & Kog, 2010).

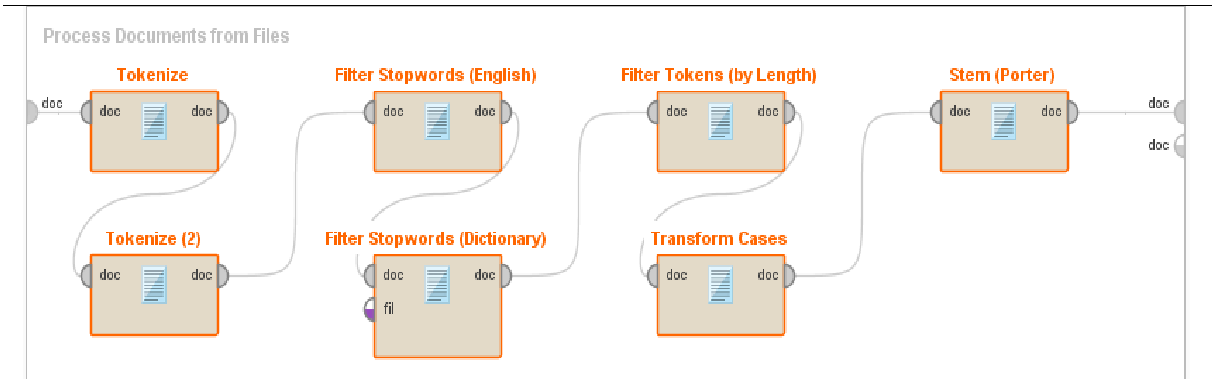

Figure 10: Preprocessing of documents in RapidMiner (Own process in Rapidminer)

After the tokenization two different stopword filters were applied, one filtering English stopwords based on a predefined dictionary and the second one filtering manually assigned stopwords. Examples of those manually defined words are names of countries, cities, companies or authors as well as other irrelevant words, like original equipment manufacturer (OEM), university or braking. They were set in an iterative fashion and excluded from the corpus. Braking or brake, as an example of such an excluded word, seems to be related to the vehicle and could potentially be of relevance for the connected vehicle in some way, at least in the first place. However, braking was not matching any of the startup descriptions and was therefore eliminated from the keyword list. Therefore, at this point, it can be said that although the keyword extraction is automated, it is still necessary to pursue some manual adjustments. The next operator filters out tokens based on their length. The minimum length was set to four characters, as suggested in RapidMiner (2016), thus the operator filters out everything that is outside that range. This is suggested as a useful step to reduce the pollution of the results (Hofmann & Chisholm, 2016). However, it may also pose the risk of losing important terms (Hofmann & Chisholm, 2016). The second last step transforms all cases of words into lower case. Finally, the Stem operator was stemming all words based on the Snowball stemming algorithm (Willett, 2006) wich iteratively replaces word suffixes according to preset rules (RapidMiner, 2016).

After the preprocessing step [5], the calculations of the TF-IDF values for each word and each document in the corpus were performed. The result of the TF-IDF calculation is a matrix showing the TF-IDF values for each word of each document and result in an array of 238 examples (documents) and 4011 attributes (identified words). An excerpt of this matrix is shown in Appendix 2.

In section [6] the clustering of the documents and thus the clustering of the respecting keywords was performed. K-means requires a manually defined number of clusters. However, there is no hard rule to define the number of clusters (RapidMiner, 2016). For that reason it was tested what might be an optimal value for the number of clusters (k), using a cluster distance performance operator as well as a comparison of the resulting keywords between different cluster sizes. The performance operator calculates the average intra-cluster distance as well as the inter-cluster distance as part of the Davies-Bouldin index (Kotu, 2014; RapidMiner, 2016). To examine the performance values for different numbers of clusters the performance was calculated for 6, 8, 10, 12, 14, 16, 18, 20, 22 and 30 clusters. The resulting distribution of items per cluster can be found in Appendix 4 and the resulting performance values are depicted in Table 10. According to the performance values, a large cluster size of e.g. k=30 yielded better values for intra- as well as inter-cluster distance. However, cluster sizes larger than 20 resulted in very small numbers of items per cluster (see Appendix 4). With a number of k=22 clusters some clusters droped below four items in a cluster and with k=30 clusters, the lowest number of items in a cluster was one. If the number of items droped below four the cluster represented a small fraction of the entire corpus (1.5%). This poses the risk of bias as keywords from these low quantity clusters are in the subsequent similarity analysis equally weighted. Thus, in the worst case the keywords might be based on a single document. Furthermore the computation of for example 30 clusters takes up to 10 minutes, making it considerably time consuming.

Measure	**Cluster k=6**	**Cluster k=8**	**Cluster k=10**	**Cluster k=12**	**Cluster k=14**	**Cluster k=16**	**Cluster k=18**	**Cluster k=20**	**Cluster k=22**	**Cluster k=30**
Average intra-cluster distance	0.864	0.847	0.829	0.817	0.804	0.788	0.778	0.766	0.753	0.705
Davies-Bouldin index	4.732	4.631	4.395	4.355	3.986	3.822	3.731	3.519	3.426	2.863

Table 10: Determination of cluster performance

Secondly, a comparison of the 200 top keywords of cluster sizes k=10 and k=20 was conducted. From the cluster (C) k=10 the top 20 keywords (K) per cluster (10C20K) and from k=20 the top ten keywords per cluster (20C10K) were extracted for the comparison. The results show that between C20K10 (173 keywords) and C10K20 (159 keywords) 92 were equal, see Table 12, p.61. Taking the results of the cluster performance and the keyword comparison a cluster number of k=20 was considered the most appropriate choice.

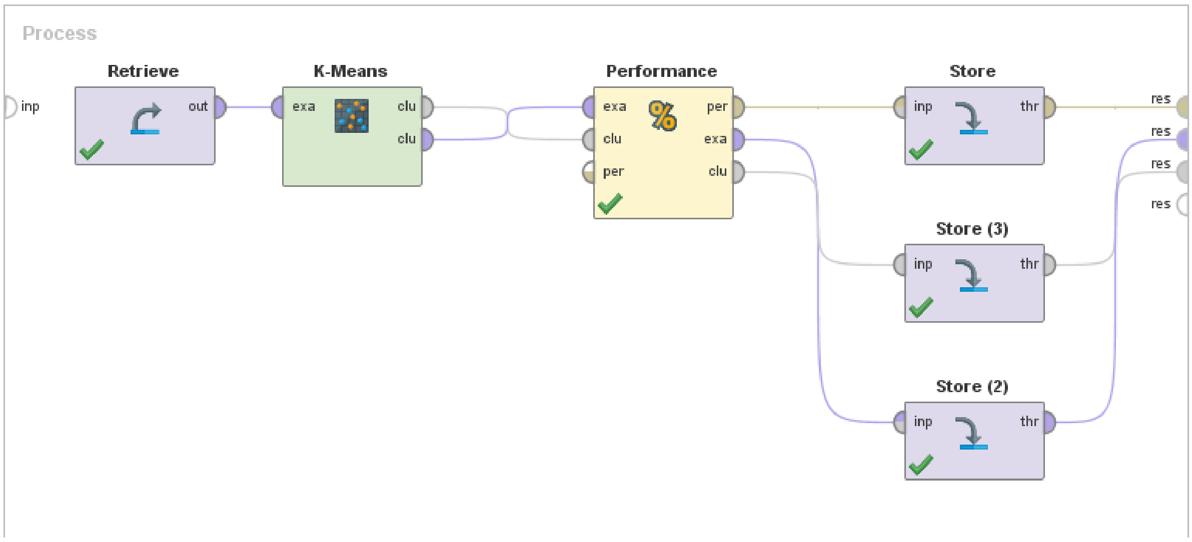

Figure 11: Clustering of documents in RapidMiner (Own process in Rapidminer)

Based on proposed evaluation of the best cluster size, the K-means clustering algorithm was set to cluster the documents into 20 clusters and perform ten runs, as proposed by default in RapidMiner (2016), of iterative cluster categorization. Due to this reason the number of clusters was set in an iterative process so that every cluster does contain a reasonable number of documents. The operators and the process of the clustering in RapidMiner are shown in Figure 11. Lastly, in section [7] the top ten keywords of each cluster of the 20 clusters were extracted and keyword duplicates were filtered out. The keywords are outlined in the results section (see Chapter 4.1)

3.3.3 Similarity analysis

The third step of the research process (process part [C]) is the similarity analysis. The similarity analysis, or measure of proximity, calculates the similarity between two records (Kotu, 2014). Within this research, the similarity was measured between the keywords and a short description of the startup.

As described earlier, the data and information stored in the organizations description was used to identify potential startups, based on the keywords extracted in the preliminary step via the text mining technique.
As not all of these companies are relevant to the connected vehicle, some filters were applied to filter out companies that have no relevance for the subsequent analysis (see Chapter 3.2). Moreover, an additional benefit is the reduction of data that has to be computed. Based on the definition of a startup (see Chapter 2.3.2), only companies younger than ten years were included (section [8]). Therefore, all company sets founded before the year 2006, as well as all data sets of companies that are no longer existent, were eliminated (section [9]). Furthermore, as some industries have none or only a very little relevance for the automotive industry certain industries were removed, such as: advertising, agriculture, biotechnology, clothing, education, events, food and beverages, healthcare, real estate, sports, beauty, coworking, craft beer, dating, delivery, dental, elder care, email, fitness, gambling, funerals, hospitality, human resources, ebooks, children, home improvement and accounting.
This prefiltering led to a total amount of 24921 relevant startups. To ensure that the ideas and innovations behind those companies have a certain maturity level, only enterprises that received more than one funding were included (section [11]). The retention of funding is perceived as accreditation and assessment of the company's idea, innovation or business model and thus serves as an indicator of maturity (Ghezzi et al., 2016). This further reduced the number of startups that go into the similarity calculation to 8007.
In order to identify potential startups in the domain of the connected vehicle, the utilization of the extracted keywords as a representation of the connected vehicle and the scanning of the Crunchbase database for potential startups addressing relevant topics for the connected vehicle was proposed. Therefore, the calculation of a similarity index between keywords and startup description was conducted. It is suggested that the index serves as an indication for startups that could potentially be a valuable source of innovation in the connected vehicle domain, maybe even of a disruptive nature. The index was calculated by determining the similarity between the startup description and the keyword. If the keyword appeared in the description, the digit 1 and if not the digit 0 was assigned to the respective keyword. After the similarity for each keyword is calculated, the index for every company was derived. An excerpt of the resulting similarity matrix can be seen in Appendix 3.
Lastly, for structuring purposes the identified startups have been coded manually using open coding via '*in vivo*' codes combined with an axial coding to further combine the identified categories as suggested for a grounded theory methodology (Saunders et al., 2011). Subsequently and as outlined in Chapter 3.1 the results were visualized in a startup map.

3.4 Quality of the research

For ensuring a high quality of the research, certain aspects to enhance the overall quality of the research were considered. As Cooper and Schindler (2014) note, good research creates reliable results that are achieved through professionally conducted procedures. This professional procedure refers to eight steps as outlined by Cooper and Schindler (2014). (1) The purpose of the research has to be clearly defined. (2) The research process needs to be described in detail. (3) The design of the research has to be planned thoroughly. (4) The research has to meet high ethical standards. (5) The limitations of the research need to be revealed. (6) The analysis for decision makers has to be adequate. (7) The findings must be presented in an unambiguous manner. (8) The researcher has to provide experiences and credentials with its research.

Furthermore, Saunders et al. (2011) outline reliability as well as validity as decisive criteria for good research results. Reliability refers to the extent the research procedure can deliver consistent findings (Saunders et al., 2011). Validity, on the other hand, addresses the issue of whether the findings are, what they seem to appear as (Saunders et al., 2011). Moreover, what Saunders et al. (2011) mention under the term external validity, also called generalizability of the results, addresses the question to what extend the research results or procedures apply to other research contexts.

The research adheres to propositions made by Cooper and Schindler (2014) as well as Saunders et al. (2011) to achieve a high quality of the research results.

The eight quality criteria provided by Cooper and Schindler (2014) were considered throughout the whole thesis and mostly have a dedicated chapter addressing the underlying quality requirements. The quality criteria provided by Saunders et al. (2011) are rather implicitly to the entire research process and the results. To ensure reliability Saunders et al. (2011) outline four potential threats that need to be taken into account, the subject or participant error, subject or participant bias, and the observer bias or error. The author tried to overcome these threats to bias and error by establishing a high level of transparency, to make the research process reproducible (Bush, 2012). The validity of the results was achieved by mining a large amount of text documents on the one hand and the triangulation via different sources (professional journals books) of these documents on the other hand. Triangulation is a useful method to ensure validity, which was proposed by Bush (2012).

4 Results

The following chapter presents the research results. These include the keywords that were extracted from professional literature as well as the startups which were identified from the Crunchbase database, using the extracted keywords.

4.1 Keyword extraction

The extraction of the keywords from the professional as well as the scientific literature on the topic of the connected vehicle was performed in accordance with Chapter 3.3.1.
The calculation of the TF-IDF values for all words and documents eventuated in a matrix with 238 documents and 3958 attributes. Due to the size of this matrix, only an excerpt can be shown (see Table 11). The TF-IDF values were only an interim result and were further processed in the clustering of the documents via the K-means clustering approach.

Row	File name	complex	display	driver	electron	radar	respect
1	File 1	0.0000	0.0000	0.0036	0.0000	0.0000	0.0000
2	File 2	0.0000	0.0464	0.0272	0.0000	0.0000	0.0123
3	File 3	0.0062	0.0124	0.0410	0.0000	0.0703	0.0000
4	File 4	0.0321	0.0000	0.0062	0.0066	0.0000	0.0591
5	File 5	0.0000	0.0756	0.0118	0.0126	0.0000	0.0080
6	File 6	0.0000	0.0000	0.0286	0.0304	0.0000	0.0097
7	File 7	0.0067	0.0000	0.0026	0.0056	0.0000	0.0000
8	File 8	0.0064	0.0000	0.0175	0.0000	0.0000	0.0068
9	File 9	0.0000	0.1278	0.0783	0.0379	0.0000	0.0000
10	File 10	0.0000	0.0371	0.0289	0.0154	0.0000	0.0098
11	File 11	0.0000	0.0000	0.0010	0.0061	0.0000	0.0418
12	File 12	0.0000	0.0000	0.0000	0.0015	0.0000	0.0205
13	File 13	0.0000	0.0000	0.0033	0.0000	0.0000	0.0140
14	File 14	0.0020	0.0000	0.0000	0.0000	0.0000	0.0041
15	File 15	0.0046	0.0000	0.0243	0.0057	0.0000	0.0000
16	File 16	0.0000	0.0000	0.0006	0.0014	0.0000	0.0052
17	File 17	0.0029	0.0000	0.0000	0.0048	0.0000	0.0276
18	File 18	0.0028	0.0000	0.0000	0.0070	0.0000	0.0237
19	File 19	0.0000	0.0113	0.0000	0.0012	0.0000	0.0030
20	File 20	0.0000	0.0000	0.0000	0.0033	0.0000	0.0126

Table 11: Keyword extraction TF-IDF matrix (Keywords randomly selected)

The K-means clustering again resulted in a matrix. All the attributes were grouped into twenty clusters, which were set manually in K-means clustering. Subsequently, the top ten keywords per cluster were derived and aggregated into one keyword list, resulting in a total amount of 200 keywords. In the next step, the 200 keywords were filtered for duplicates, which were removed, leading to a total of 173 unique keywords (see Table 12).

Extracted keyword set C20K10							
acceler	carshar	delivery	flexray	leon	powertrain	screen	stop
adopt	cell	density	freeway	linux	predict	seat	sync
advertis	channel	departur	frequenc	local	premium	secur	taxi
algorithm	chapter	deploy	gnss	messag	primit	semiconductor	telemat
analy	charg	diagnost	govern	metric	privacy	sensor	testb
antenna	chip	digit	grid	modul	privat	session	threat
atom	city	digitalis	gridlock	motorist	processor	share	ticket
attack	cloud	display	hadoop	multi	profit	sign	touch
audio	cluster	disrupt	hail	navig	protect	simul	tour
authent	collis	domain	har	node	protocol	site	traffic
autom	compass	driver	hardwar	obstacl	radio	situat	transport
autonom	concept	ecosystem	hazard	organis	railway	slide	undergo
bandwidth	congest	electr	headway	overhead	reilhac	slip	unif
battery	connector	electron	infotain	packet	relay	smart	unman
beacon	constitu	emiss	intel	park	request	social	valid
bev	contact	energy	interior	patent	retriev	song	video
buse	context	enterpris	intersect	pattern	revenu	space	vissim
busi	curv	estim	intuit	peer	robot	spectrum	voltag
buyer	cyber	ethernet	junction	peopl	rout	stake	warn
camera	data	filter	lake	planner	safeguard	stamp	
campaign	databas	firmwar	lane	policy	sale	standstil	
car	delivery	flexray	learn	powertrain	scatter	stop	

underlined keywords indicate similarity between C10K20 and C20K10 (98 of 173 are equal)

Table 12: Keywords from cluster (C) k=20 and top ten keywords (K) per cluster

4.2 Identification of potential startups

This chapter provides the results of the identified startups. The generated keywords from the preliminary step of the research process and the 8007 (see Table 13) startup descriptions from the Crunchbase database were utilized to measure their textual similarity.

Based on the similarity between each keyword and the company's description a similarity index was calculated for every startup. The similarity index represents the number of keywords that were found to be similar to the company's description.

In total, the similarity analysis identified 5494 startups that have at least a similarity index of one and thus at least one match between the keywords and the startup description.

By implication, this means that the similarity analysis could only reduce the original data set by 2513 startups.

In order to identify the relevant startups for each similarity index, all startups remaining in the database were coded manually, resulting in 319 startups identified as relevant with regard to the connected vehicle. The startups were grouped into ten main categories as visualized in Figure 12. The ratios calculated in Table 13 showed that the higher the similarity index, the more likely are the according startups to be relevant in the area of the connected vehicle. This can be seen in the ratio calculated from the number of startups after coding divided by the number of startups before coding (see Table 13). For a similarity index of one and two, the relevance of startups towards the topic connected vehicle is relatively small with a ratio of 1.7% and 3.6%. Nevertheless, those startups still account for 106 (47+59) out of 319 startups.

Similarity index	Number of startups before coding	Number of startups after coding	Ratio of number of startups after coding/number of startups before coding	Ratio of startups after coding /total number of startups after coding per similarity index
0	2513	0	-	-
1	2755	47	1.7%	14.7%
2	1619	59	3.6%	18.5%
3	763	111	14.5%	34.8%
4	272	66	24.3%	20.7%
5	67	26	38.8%	8.2%
6	15	8	53.3%	2.5%
7	1	0	0.0%	0.0%
8	2	2	100.0%	0.6%
∑	8007	319		

Table 13: Similarity index performance

The similarity analysis proved that it was able to reduce the number of startups in the database. Thus, it was also able to reduce the required manual effort, as it was not necessary to manually scan all startups in the database. However, a manual coding and evaluation is still necessary. Overall, the method proofed to be valuable and beneficial as 319 startups could be identified as relevant with regard to the connected vehicle. Those startups are visualized in a startup map (see Figure 12).

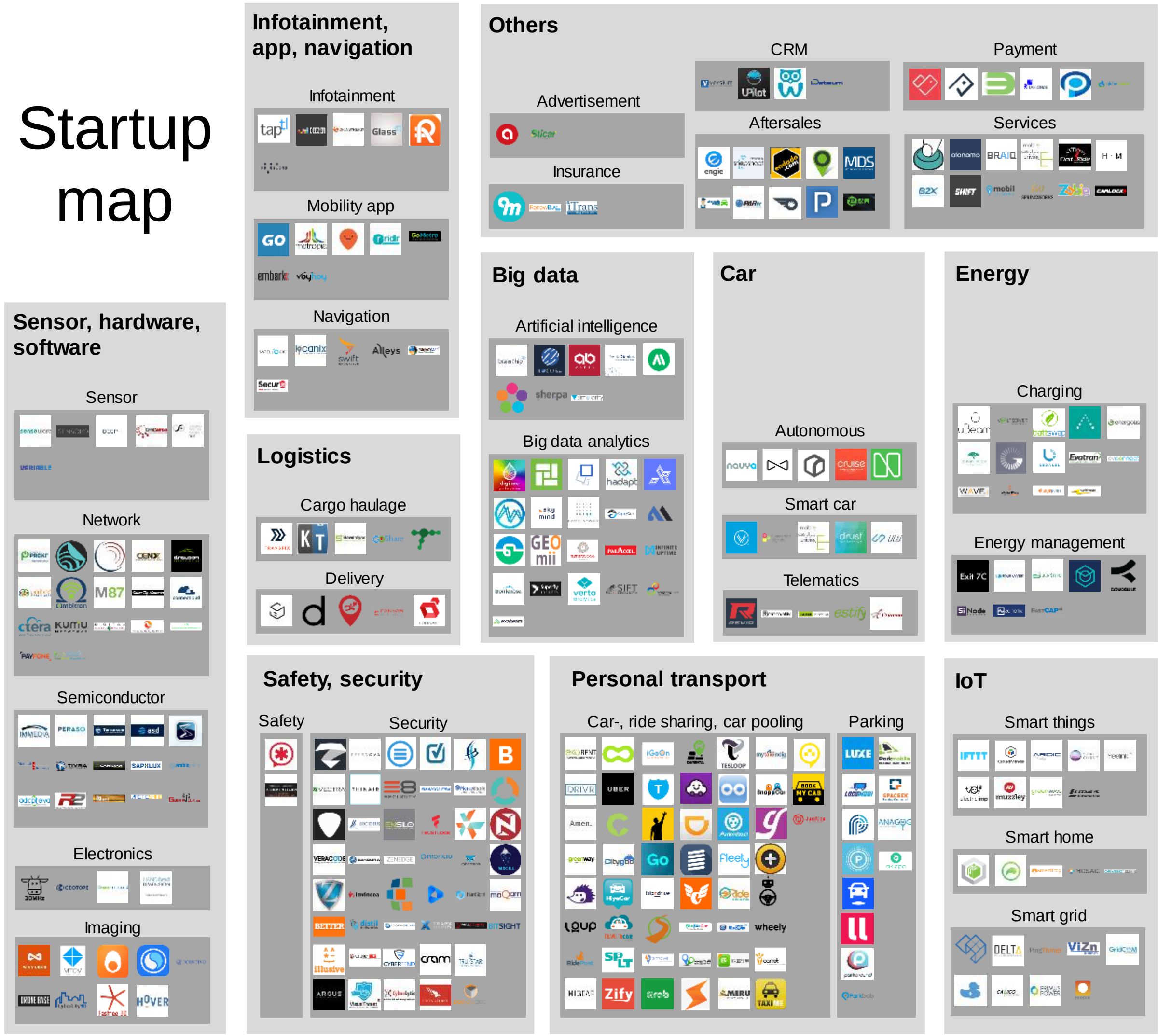

Figure 12: Startup map - manually coded startups resulting from the similarity analysis (Own illustration)

5 Conclusion

This chapter provides a conclusion of the theoretical findings, integrates them with the proposed methodological framework and provides an answer to the research questions in Chapter 1.2. Furthermore, it is described how the results contribute to theory and business practice. Additionally, research limitations are outlined and opportunities for future research are presented.

5.1 Summary

This chapter provides a summary of the conclusions drawn from each chapter of the literature review. Furthermore, it presents a summarization of the research results. Concerning the literature review it outlines the contribution of each stream of the literature to corporate foresight as well as the role or contribution of startups within each perspective for corporate foresight. Furthermore, the theoretical findings from the literature analysis in the field of innovations and startups are summarized. Moreover, this chapter will connect the theoretical underpinnings, the proposed methodological approach, and the research results. It aims to answer both research sub-questions and thereby the main research question.

Within the strategic management perspective, the main focus lies on the scanning for and gaining of information about the company's environment as well as on sensing signals of change. Pioneers or special groups are considered as a valuable information source within that perspective, and thus startups are considered a source to sense the signals of change accordingly. Specific methods to sense and utilize these environmental information and indications of change for corporate foresight activities are addressed and provided within the future research perspective. Looking more closely into how these information are transformed into innovation is addressed in the innovation management perspective. It is concerned on how innovation can be fostered and achieved successfully. Startups play a decisive role in this context as they possess particular benefits regarding their organizational culture and structure as well as their knowledge. The causation and effectuation perspective represents two distinct logics on how to create and discover opportunities or innovation and contributes to corporate foresight by making the corporate entity aware of these distinct logics. Startups are relevant from the view of this perspective as they are recognized as showing more effectual logic which is beneficial for innovation, specifically for those innovations that are of a radical or disruptive nature. Taking a closer look into capabilities, it can be concluded that corporate foresight itself can be regarded as a dynamic capability. Corporate foresight allows a company to renew in accordance with the environmental conditions. Startups might serve as valuable sources within this perspective as startups are considered to have different, expedient and superior skills in sensing, seizing and reconfiguring dynamic capabilities. Corporate entrepreneurship is considered to contribute to corporate foresight as it deals with the recognition and identification of opportunities and thus with the creation of innovations. This perspective shows that startups should be considered as a valuable external source of opportunities or innovations for incumbent firms. The historical development of corporate foresight illustrates the growing relevance and sophistication of this research stream. The recent developments show that foresight has been utilizing ICT-Tools to a greater extent and that it is expected to increase, allowing a more automated information retrieval and sense making. This line of argumentation is supported by a new stream within the academic literature. This new stream is concerned with methodological approaches on how opportunities, technologies or innovation can be discovered, utilizing ICT-Tools and advanced analytics. Furthermore, the stream shows that a variety of data sources can serve to identify opportunities, technologies or innovations. However, this specific stream of the literature is not connected to the traditional corporate foresight literature yet.

By expanding the view and by looking specifically into potential sources of innovations, the literature provides an ample list of possible innovation sources. However, startups are not explicitly mentioned as a potential source within that stream of the literature. On the contrary, the startup literature validated that startups have benefits compared to incumbent firms. Thus, cooperation between startups and incumbents are highly valuable as a source for corporate innovations and the internalization of external innovations. The developments of the last decade within the research stream of corporate venturing and open innovation show that startups are increasingly recognized as a source of external innovations. Summarized, startups are considered a valuable input within the different corporate foresight perspectives, due to environmental information, specific characteristics like the organizational structure, the mental logic of decision making and action taking or the dynamic capabilities. Furthermore, it can be concluded that parts of academia and business practice are already considering startups as a valuable source for innovations and that large corporate entities are already engaging with startups. However, startups are not yet considered as a specific source for environmental information or innovations within corporate foresight. Table 14 provides an overview of the conclusions drawn from the literature review.

	Research stream	Conclusion
1	Strategic management perspective	The strategic management systems, environmental scanning as well as different sources for environmental information build the foundation of the corporate foresight research. Predominantly important within this research stream is the scanning of the corporate environment for information about change. Thereby pioneers, specialized groups and thus startups are considered a relevant source for information.
2	Innovation management perspective	The innovation management perspective contributes to corporate foresight by outlining the process and requirements for a successful innovation management. Specifically, how innovative ideas occur and how they can be fostered. Concerning startups, it shows that they possess particular benefits regarding organizational culture, organizational structure as well as knowledge.
3	Future(s) research perspective	The futures research perspective is relevant for corporate foresight as it provides the methods that are necessary for its conduction. The methods do thereby vary, as there are no universal success factors and the methods need to be chosen dependent on the respecting environmental conditions.
4	Dynamic capability perspective	The dynamic capability perspective contributes to the corporate foresight literature as corporate foresight can be regarded a dynamic capability as it enables a company to renew its resources in a dynamic market environment. Looking into startup context, it can be concluded that startups are valuable for incumbents as new ventures have different, expedient and superior skills in sensing, seizing and reconfiguring dynamic capabilities.
5	Causation and effectuation perspective	The causation and effectuation perspective contributes to corporate foresight as it describes two distinct logics of how opportunities and innovations are created. Concerning startups, it can be concluded that they tend to be more effectual than incumbents and thus show a higher likelihood of developing innovations specifically radical and disruptive ones.
6	Corporate entrepreneurship perspective	The research stream of corporate entrepreneurship deals with the creation, recognition and identification of opportunities and their contribution to the corporate foresight literature by examining how a corporation can create innovations. Concerning startups, it can be concluded that the research stream of corporate entrepreneurship specifically points out that the opportunity creation can be internal or external to the company and that external ventures should be considered a valuable source for opportunities.
7	Development of corporate foresight	Recent corporate foresight developments in utilizing ICT-Tools to support and complement traditional corporate foresight activities advance more and more into an automated information retrieval and sense making. Future corporate foresight activities will most likely expand into this domain and use a variety of data sources to identify relevant environmental information, opportunities and innovation.
8	Methodological approaches for opportunity, technology and innovation discovery	This rather new field of research about specific methodological approaches is not yet connected to the field of corporate foresight, although corporate foresight is requesting more advanced methodological approaches to enhance the corporate foresight performance. Furthermore, it shows that a variety of data sources can serve to identify potential opportunities, technologies or innovations and support the corporate foresight process.
9	Potential sources of innovation	The academic literature on potential sources of innovation provides an extensive overview of potential sources. However, startups and entrepreneurial firms are not explicitly considered as a potential source in the academic literature yet, despite their benefits and superiority in many aspects relevant for a company's innovativeness.
10	Startups as a potential source for corporate innovation	Based on the outlined benefits of startups and an engagement between startups and established corporations it should be a prerequisite for corporations to incorporate startups into their corporate foresight activities. Incumbents who are in search of innovations, specifically those of a radical or disruptive nature, should actively engage with startups, utilize their advantages and internalize external innovation. This is supported in the recent developments in the field of corporate venturing and open innovation as startups are already considered a highly relevant source for corporate innovations there.

Table 14: Conclusions from the literature review

Having outlined that corporate foresight should integrate startups more actively into its activities due to the benefits for a corporate firm, now the second sub-research question will be answered. The second sub-research question asked how startups can be identified utilizing a startup database? Chapter 3 outlined a structured methodological approach how startups can be identified from a startup database utilizing text mining. The approach consists of three sections. In a first step expert literature on the connected vehicle is identified from scientific and professional sources. The second step utilized the expert literature and an automated text mining technique to extract keywords which are characteristic for the connected vehicle. In the third step these keywords were employed to identify relevant startups from the database.
The results in Chapter 4.2 showed that a startup database is a useful tool to identify startups as 319 startups relevant to the connected vehicle could be identified with the proposed approach. However, even the proposed approach requires some additional manual effort to derive the startups from the database. The visualization of the startups in the startup map provide a comprehensive overview of startup activities regarding the connected vehicle (see Figure 12). A comparison of incumbents' cooperation and investment activities with the identified startups shows that many of them are already investing in or cooperating with startups identified from the database, like BMW AG with investments in Chargepoint, Moovit, and Parkmobile (BMW AG, 2016a), the cooperation between BMW and IFTTT (BMW AG, 2016b) or the strategic partnership between Volkswagen and Gett (Volkswagen AG, 2016). The proposed approach could have provided a useful tool to support the assessment of potential investment alternatives, scan the competitive environment and identify companies providing related technologies, products or services.

5.2 Contribution to theory and business practice

The aim of the research was to show how corporations can utilize startups as a source of innovations within corporate foresight. In particular, the research anticipated to answer what role startups play as a potential source for innovations and how they can be identified utilizing a startup database. The underlying motivation was induced by the identified gaps in the corporate foresight literature. Firstly, startups were not yet specifically considered as a relevant source for environmental information or innovations in the corporate foresight literature (see Chapter 2.2 and 5.1). Secondly, several scholars within the corporate foresight literature outlined that corporate foresight needs more advanced and automated foresight methodologies and tools (Heger & Rohrbeck, 2012; Heuschneider & Herstatt, 2016; Paliokaitė & Pačėsa, 2015; Rohrbeck, Battistella et al., 2015).

Taking the identified gap in the literature and the research questions as a guiding framework, the thesis addresses three general topics, (1) corporate foresight, (2) corporate venturing, innovation, and startups as well as (3) opportunity-, idea- and innovation discovery.

On a theoretical basis, this thesis examines that startups should be incorporated into the corporate foresight activities of a firm. Firstly, it investigates from the view of different corporate foresight perspectives how they contribute to corporate foresight and how startups might contribute within each perspective. Secondly, as innovations are inherently important for corporate foresight, potential sources of innovation, as well as environmental information, are investigated in general, and more specifically the role of startups as a valuable source for innovations is examined on a theoretical basis. Thirdly, particular advantages and drawbacks of startups over incumbent firms and the benefits of a collaboration between the both are outlined, describing in detail what makes startups valuable for incumbent firms. Fourthly, the scientific work of a specific literature stream on the opportunity, technology, and innovation discovery is examined, outlining different methodological approaches on how opportunities, technologies, or innovations can be detected. Theoretically, the thesis contributes to theory and practice by closing the identified gap between the different research streams from a theoretical perspective and outlines that there is a need to incorporate startups more actively into the corporate foresight activities. Moreover, it shows the requirement for a more thorough methodological approach to identify startups.

The present thesis contributes particularly to two existing gaps in the scientific literature. The first gap exists between the corporate foresight research and the research stream on the different methodological approaches to opportunity, technology, and innovation discovery. So far there is no connection between both research streams, although the stream on opportunity identification clearly addresses important aspects of the foresight literature (Seo et al., 2016; Yoon et al., 2015; Yoon et al., 2016; Zhang et al., 2016) and the foresight literature is requesting more sophisticated methods to identify opportunities, technologies, innovations or environmental information (Heger & Rohrbeck, 2012; Heuschneider & Herstatt, 2016; Paliokaitė & Pačėsa, 2015; Rohrbeck, Battistella et al., 2015).

The second gap exists between the foresight literature and the literature on potential sources of innovation as well as startups as a specific but not well-recognized source for environmental information or as a source for the internalization of external innovations. Thus, it enables foresight researchers and practitioners to gain insights into the valuable potential of startups.

On a practical basis, this thesis contributes by proposing a methodological framework or process to identify startups from a startup database. The framework rests on the methodological approaches from the opportunity, technology, and innovation discovery literature. Several research contributions of that stream proposed to utilize text mining as a methodological approach for the detection of environmental information, technological opportunities or innovations (Seo et al., 2016; Thorleuchter & van den Poel, 2015; Yoon et al., 2016; Zhang et al., 2016). Within this report, the existing approaches were combined and extended to identify startups from a startup database.

The proposed methodological approach enables practitioners to enhance their corporate foresight activities. It empowers them to identify startups in a particular area of interest and thus facilitates the attainment of environmental information or ideas which could function as a trigger and be further developed by research and development as well as design units. This approach could contribute to the identification of product, technology or business model innovations which would most probably not be generated and fostered by the existing research and development or product design units of the firm. Thereby the number of required experts and the costs for identification of potential startup activities are reduced. This enables foresight researchers or practitioners to quickly examine different fields of interest. Lastly, the startup map, provides practitioners an overview of startups in the field of the connected vehicle.

More generally, the theoretically derived knowledge and the outlined methodological approach enhance the overall value contribution of corporate foresight as outlined by Rohrbeck and Schwarz (2013) as well as Paliokaitė and Pačėsa (2015). The approach supports the capability to enlarge the view on the corporate environment, reduces uncertainty and enhances the ability to conduct corporate exploration activities.

Besides the contribution to corporate foresight, the results could additionally be beneficial within corporate venture units as they also need to identify inter alia startups in order to recognize and explore new business opportunities (Hill & Birkinshaw, 2014).

5.3 Limitations and future research

The conducted research also comprises some limitations, especially because of its exploratory nature and the application of advanced data processing and analytics to a research field that is characterized by expert opinions.

A key assumption of the present research is that relevant keywords can be extracted from professional journal articles and books. Although several different sources and a large amount of literature were utilized, the potential that a different literature corpus might extract different keywords still exists. Nevertheless, this is mitigated to some extent by the clustering of the documents. However, future research should investigate whether a corpus derived from different sources extracts significantly different keywords. The sources could include other professional or scientific literature, news articles as utilized by Iglesias et al. (2016) or even twitter data as suggested by Kayser and Blind (2016).

The coding of the identified startups into categories and thus the identification of the 319 startups in the area of the connected vehicle might be biased due to the evaluation by a single researcher. In order to avoid a manual involvement of the researchers or of experts, future research should further improve the automated extraction of keywords as well as the measurement of similarity between the extracted keywords and the startup description. A first possibility to improve the similarity analysis could be the usage of a longer startup description. This would yield more information about the startup and could thus deliver better results from the similarity analysis as there is a higher probability that the extracted keywords match the description. A second possibility that should be examined by future research is whether a weighting of the extracted keywords improves the performance of the similarity analysis and reduces the necessary manual effort. Currently, the similarity analysis only checks whether the keyword occurs in the description, but by identifying more or less important keywords, the performance of the similarity analysis could potentially be improved. A third area of future research could be the examination of the performance of different cluster algorithms. With regard to clustering algorithms other researchers propose the necessity of their development and refinement in future research (Zhang et al., 2016).

The multi-dimensional theoretical analysis of the role and value of startups for corporate foresight shows that startups could be a valuable source of environmental information and innovation. Thus, startups can be useful within corporate foresight. However, a thorough evaluation of the proposed methodological approach within a practical foresight project on a longitudinal basis is necessary to proof the suggested value of startups within corporate foresight. The evaluation framework for future studies as proposed by Piirainen et al. (2012) could serve as a basis for such an assessment. Future research should not only investigate which value startups can bring to corporate foresight, it should also be assessed how valuable or beneficial a specific startup could be for the corporate entity. This could include for example the corporate fit or the potential value of the startup for the incumbent. Similarly to the future research suggestion of Yoon et al. (2016) it would be beneficial to incorporate the proposed approach into a software to make its application more convenient.

References

1. ABDC. (2016). ABDC Journal Quality List · Australian Business Deans Council. Retrieved from http://www.abdc.edu.au/pages/abdc-journal-quality-list-2013.html
2. Abernathy, W. J., & Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction. *Research Policy*, *14*(1), 1.
3. ABS. (2015). ABS Academic Journal Guide 2015. Retrieved from http://www.du-bisheng.com/www/docs/ABS2015.pdf
4. Ackoff, R. (1970). A concept of corporate planning. *Long range planning*, *3*(1), 2–8.
5. Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews*, *8*(1), 21–47.
6. Aguilar, F. J. (1967). *Scanning the business environment: Studies of the modern corporation* (1. print). New York: Arkville Press.
7. Ahuja, G., & Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, *22*(6-7), 521–543.
8. Albright, K. S. (2004). Environmental scanning: radar for success: environmental scanning--the internal communication of external information about issues that may influence an organization's decision-making process--can identify emerging issues, situations, and potential pitfalls that may affect an organization's future. *Information Management Journal*, *38*(3), 38.
9. Alvarez, S. A., & Barney, J. B. (2001). How entrepreneurial firms can benefit from alliances with large partners. *The Academy of Management Executive*, *15*(1), 139–148.
10. Alvarez, S. A., & Barney, J. B. (2007). Discovery and creation: Alternative theories of entrepreneurial action. *Strategic Entrepreneurship Journal*, *1*(1-2), 11–26.
11. Ambrosini, V., & Bowman, C. (2009). What are dynamic capabilities and are they a useful construct in strategic management? *International Journal of Management Reviews*, *11*(1), 29–49.
12. Amsteus, M. (2012). The Origin of Foresight. *World Futures*, *68*(6), 390–405.
13. Ansoff, H. I. (1975). Managing strategic surprise by response to weak signals. *California Management Review*, *18*(2), 21–33.
14. Ansoff, H. I. (1980). Strategic issue management. *Strategic Management Journal*, *1*(2), 131–148.
15. Ansoff, H., Avner, J., Brandenburg, R. G., Portner, F. E., & Radosevich, R. (1970). Does planning pay?: The effect of planning on success of acquisitions in American firms. *Long range planning*, *3*(2), 2–7.
16. Antoncic, B., & Hisrich, R. D. (2001). Intrapreneurship. *Journal of Business Venturing*, *16*(5), 495–527.
17. Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity identification and development. *Journal of Business Venturing*, *18*(1), 105–123.
18. Arend, R. J. (2014). Entrepreneurship and dynamic capabilities: How firm age and size affect the ‘capability enhancement–SME performance’ relationship. *Small Business Economics*, *42*(1), 33–57.
19. Assink, M. (2006). Inhibitors of disruptive innovation capability: A conceptual model. *European Journal of Innovation Management*, *9*(2), 215–233.
20. Baark, E., Antonio, K. W., Lo, W., & Sharif, N. (2011). Innovation sources, capabilities and competitiveness: evidence from Hong Kong firms. *DIMA Dynamics of Institutions and Markets in Europe - Conference*.
21. Baregheh, A., Rowley, J., & Sambrook, S. (2009). Towards a multidisciplinary definition of innovation. *Management Decision*, *47*(8), 1323–1339.
22. Barney, J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of management*, *27*(6), 625–641.
23. Barney, J. B. (1995). Looking inside for competitive advantage. *Academy of Management Perspectives*, *9*(4), 49–61.
24. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, *17*(1), 99–120.
25. Barney, J. B. (1986). Organizational culture: Can it be a source of sustained advantage? *Academy of Management Review*, *11*(3), 656.
26. Baron, R. A. (2006). Opportunity Recognition as Pattern Recognition: How Entrepreneurs "Connect the Dots" to Identify New Business Opportunities. *Academy of Management Perspectives*, *20*(1), 104–119.
27. Baskarada, S., Shrimpton, D., Ng, S., Cox, D., & Saritas, O. (2016). Learning through foresight. *Foresight*, *18*(4).
28. Basu, S., & Wadhwa, A. (2013). External Venturing and Discontinuous Strategic Renewal: An Options Perspective. *Journal of Product Innovation Management*, *30*(5), 956–975.
29. Baum, J. A. C., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic management journal*, *21*(3), 267–294.

30. Baumol, W. J. (2004). Entrepreneurial enterprises, large established firms and other components of the free-market growth machine. *Small business economics*, *23*(1), 9–21.
31. Beldow, R., Frese, M., Anderson, N., Erez, M., & Farr, J. (2009). A Dialectic Perspective on Innovation: Conflicting Demands, Multiple Pathways, and Ambidexterity. *Industrial and Organizational Psychology*, *2*(3), 305–337.
32. Berends, H., Jelinek, M., Reymen, I., & Stultiëns, R. (2014). Product Innovation Processes in Small Firms: Combining Entrepreneurial Effectuation and Managerial Causation. *Journal of Product Innovation Management*, *31*(3), 616–635.
33. Berry, M., & Kog, J. (2010). Text Mining: Applications and Theory. Retrieved from http://www.mohamedrabeea.com/books/book1_1165.pdf
34. Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review*, *91*(5), 64.
35. Blind, K., Cuhls, K., & Grupp, H. (1999). Current Foresight Activities in. *Technological Forecasting and Social Change*, *60*(1), 15–35.
36. Block, F., & Keller, M. R. (2009). Where do innovations come from?: Transformations in the US economy, 1970-2006. *Socio-Economic Review*, *7*(3), 459–483.
37. Blyler, M., & Coff, R. W. (2003). Dynamic capabilities, social capital, and rent appropriation: Ties that split pies. *Strategic Management Journal*, *24*(7), 677–686.
38. BMW AG. (2016a). BMW i Ventures. Retrieved from http://www.bmw.com/com/en/insights/corporation/bmwi_ventures/index.html
39. BMW AG. (2016b). What is BMW's IFTTT Service? | BMW Auto Blog. Retrieved from http://www.bmwoffreeport.com/blogs/827/bmw-begins-integrating-ifttt-service/
40. Bogers, M., & West, J. (2012). Managing Distributed Innovation: Strategic Utilization of Open and User Innovation. *Creativity and innovation management*, *21*(1), 61–75.
41. Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (Second edition). London: Sage Publications.
42. Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science*, *57*(5), 843–863.
43. Bray, O. H., & Garcia, M. L. (1997). Technology roadmapping: the integration of strategic and technology planning for competitiveness. *Portland International Conference on Management of Engineering and Technology (PICMET)*, 25–28.
44. Burgelman, R. A. (1983). Corporate Entrepreneurship and Strategic Management: Insights from a Process Study. *Management Science*, *29*(12), 1349–1364.
45. Bush, T. (2012). Authenticity in research–reliability, validity and triangulation. In A. Briggs, M. Coleman, & M. Morrison (Eds.), *Research Methods in Educational Leadership & Management.* London: Sage Publications.
46. BVIZ. (2016). Bundesverband Deutscher Innovations-, Technologie- und Gründerzentren e.V. (BVIZ) | English Information. Retrieved from http://www.innovationszentren.de/index.php?article_id=41
47. Cai, L., Guo, R., Fei, Y., & Liu, Z. (2016). Effectuation, Exploratory Learning and New Venture Performance: Evidence from China. *Journal of Small Business Management*, *00*(00), 1–16.
48. Carland, J. W., Hoy, F., Boulton, W. R., & Carland, J. A. C. (1984). Differentiating Entrepreneurs from Small Business Owners: A Conceptualization. *Academy of Management Review*, *9*(2), 354–359.
49. Cassiman, B., Di Guardo, M. C., & Valentini, G. (2010). Organizing links with science: Cooperate or contract? *Research Policy*, *39*(7), 882–892.
50. Cassiman, B., & Veugelers, R. (2006). In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science*, *52*(1), 68–82.
51. Chandler, G. N., DeTienne, D. R., McKelvie, A., & Mumford, T. V. (2011). Causation and effectuation processes: A validation study. *Journal of Business Venturing*, *26*(3), 375–390.
52. Chandy, R. K., & Tellis, G. J. (1998). Organizing for Radical Product Innovation: The Overlooked Role of Willingness to Cannibalize. *Journal of Marketing Research*, *35*(4), 474.
53. Chandy, R. K., & Tellis, G. J. (2000). The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation. *Journal of Marketing*, *64*(3), 1.
54. Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate Venture Capital, Value Creation, and Innovation. *Review of Financial Studies*, *27*(8), 2434–2473.
55. Chesbrough, H. (2007). Business model innovation: it's not just about technology anymore. *Strategy & Leadership*, *35*(6), 12–17.
56. Chesbrough, H., Vanhaverbeke, W., & West, J. (2006). *Open innovation: Researching a new paradigm*: Oxford University Press.
57. Christensen, C. (2011). *The innovator's dilemma: when new technologies cause great firms to fail*. New York: Harper Business (Original work published 1997).

58. Christensen, C., & Raynor, M. (2013). *The innovator's solution: Creating and sustaining successful growth*. Boston: Harvard Business Review Press.
59. Clausen, T., Pohjola, M., Sapprasert, K., & Verspagen, B. (2012). Innovation strategies as a source of persistent innovation. *Industrial and Corporate Change*, *21*(3), 553–585.
60. Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*(1), 128.
61. Cook, C. N., Inayatullah, S., Burgman, M. A., Sutherland, W. J., & Wintle, B. A. (2014). Strategic foresight: how planning for the unpredictable can improve environmental decision-making. *Trends in ecology & evolution*, *29*(9), 531–541.
62. Cooper, D. R., & Schindler, P. S. (2014). *Business research methods* (12. edition). *The McGraw-Hill/Irwin series in operations and decision sciences Business statistics*. New York: McGraw-Hill Irwin.
63. Corbett, A. C. (2005). Experiential Learning Within the Process of Opportunity Identification and Exploitation. *Entrepreneurship Theory and Practice*, *29*(4), 473–491.
64. Cornish, E. (2003). The wild cards in our future: anticipating wild cards--unexpected, life-changing events--is important in preparing for the future and coping with surprise. *The Futurist*, *37*(4), 18.
65. Corso, M., & Pellegrini, L. (2007). Continuous and Discontinuous Innovation: Overcoming the Innovator Dilemma. *Creativity and Innovation Management*, *16*(4), 333–347.
66. Crockett, D. R., McGee, J. E., & Payne, G. T. (2013). Employing New Business Divisions to Exploit Disruptive Innovations: The Interplay between Characteristics of the Corporation and Those of the Venture Management Team. *Journal of Product Innovation Management*, *30*(5), 856–879.
67. Crossan, M. M., & Apaydin, M. (2010). A Multi-Dimensional Framework of Organizational Innovation: A Systematic Review of the Literature. *Journal Of Management Studies*, *47*(6), 1154–1191.
68. Crunchbase. (2016). CrunchBase accelerates innovation by bringing together data on companies and the people behind them. Retrieved from https://www.crunchbase.com/#/home/index
69. Cuhls, K. (2003). From forecasting to foresight processes—new participative foresight activities in Germany. *Journal of forecasting*, *22*(2-3), 93–111.
70. D'Haen, J., van den Poel, D., & Thorleuchter, D. (2013). Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique. *Expert Systems with Applications*, *40*(6), 2007–2012.
71. Daft, R. L., & Weick, K. E. (1984). Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review*, *9*(2), 284–295.
72. Daft, R. L., Parks, D., & Sormunen, J. (1988). Chief executive scanning, environmental characteristics, and company performance: An empirical study. *Strategic management journal*, *9*(2), 123–139.
73. Dagnino, G. B., Levanti, G., Minà, A., & Picone, P. M. (2015). Interorganizational network and innovation: A bibliometric study and proposed research agenda. *Journal of Business & Industrial Marketing*, *30*(3/4), 354–377.
74. Daheim, C., & Uerz, G. (2008). Corporate foresight in Europe: From trend based logics to open foresight. *Technology Analysis & Strategic Management*, *20*(3), 321–336.
75. Daimler AG. (2015). Mercedes me connect. Retrieved from https://www.mercedes-benz.com/de/mercedes-me/konnektivitaet/
76. Damanpour, F. (1991). Organizational innovation: a meta-analysis of effects of determinants and moderators. (includes appendix). *Academy of Management Journal*, *34*(3), 555.
77. Damanpour, F., & Gopalakrishnan, S. (1998). Theories of organizational structure and innovation adoption: the role of environmental change. *Journal of Engineering and Technology Management*, *15*(1), 1–24.
78. Darroch, J. (2005). Knowledge management, innovation and firm performance. *Journal of Knowledge Management*, *9*(3), 101–115.
79. Dasgupta, S., & Sanyal, D. (2009). Bridge to the future: Connect your strategies in an interconnected world. *Foresight*, *11*(1), 81–93.
80. Day, G. S., & Schoemaker, P. (2004a). Driving Through the Fog: Managing at the Edge. *Long range planning*, *37*(2), 127–142.
81. Day, G. S., & Schoemaker, P. (2004b). Peripheral Vision: Sensing and Acting on Weak Signals. *Long range planning*, *37*(2), 117–121.
82. Dee, N., Gill, D., Weinberg, C., & McTavish, S. (2015). *What's the difference?* Retrieved from http://www.nesta.org.uk/sites/default/files/whats_the_diff_wv.pdf
83. Demirel, P., & Mazzucato, M. (2012). Innovation and Firm Growth: Is R&D Worth It? *Industry & Innovation*, *19*(1), 45–62.
84. DeTienne, D. R., & Chandler, G. N. (2004). Opportunity Identification and Its Role in the Entrepreneurial Classroom: A Pedagogical Approach and Empirical Test. *Academy of Management Learning & Education*, *3*(3), 242–257.

85. Dewar, R. D., & Dutton, J. E. (1986). The Adoption of Radical and Incremental Innovations: An Empirical Analysis. *Management Science*, *32*(11), 1422–1433.
86. Díaz-Díaz, N. L., & Saá Pérez, P. de. (2014). The interaction between external and internal knowledge sources: An open innovation view. *Journal of Knowledge Management*, *18*(2), 430–446.
87. Donaldson, L. (1999). The normal science of structural contingency theory. In S. R. Clegg, S. Clegg, & C. Hardy (Eds.), *Studying organization: theory and method* (pp. 51–70). London: Sage Publications.
88. Donaldson, L. (2001). *The contingency theory of organizations*. *Foundations for organizational science*. Thousand Oaks: Sage Publications.
89. Donaldson, L. (2006). The Contingency Theory of Organizational Design: Challenges and Opportunities. In R. M. Burton, D. D. Håkonsson, B. Eriksen, & C. C. Snow (Eds.), *Organization Design: The evolving state-of-the-art* (pp. 19–40). Boston: Springer.
90. Döring, N., & Bortz, J. (2016). Forschungsstand und theoretischer Hintergrund. In N. Döring & J. Bortz (Eds.), *Springer-Lehrbuch. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5th ed., pp. 157–179). Berlin, Heidelberg: Springer.
91. Dosi, G. (1982). Technological paradigms and technological trajectories. *Research Policy*, *11*(3), 147–162.
92. Drucker, P. F. (1998). The discipline of innovation. *Harvard Business Review*, *76*(6), 149.
93. Drucker, P. F. (1959). Long-Range Planning: Challenge to Management Science. *Management Science*, *5*(3), 238–249.
94. Dunlap-Hinkler, D., Kotabe, M., & Mudambi, R. (2010). A story of breakthrough versus incremental innovation: Corporate entrepreneurship in the global pharmaceutical industry. *Strategic Entrepreneurship Journal*, *4*(2), 106–127.
95. Durst, C., Durst, M., Kolonko, T., Neef, A., & Greif, F. (2015). A holistic approach to strategic foresight: A foresight support system for the German Federal Armed Forces. *Technological Forecasting and Social Change*, *97*, 91–104.
96. Dushnitsky, G., & Shaver, J. M. (2009). Limitations to interorganizational knowledge acquisition: The paradox of corporate venture capital. *Strategic Management Journal*, *30*(10), 1045–1064.
97. Dutton, J. E., & Ottensmeyer, E. (1987). Strategic Issue Management Systems: Forms, Functions, and Contexts. *The Academy of Management Review*, *12*(2), 355–365.
98. Dyer, J. H., Gregersen, H. B., & Christensen, C. (2008). Entrepreneur behaviors, opportunity recognition, and the origins of innovative ventures. *Strategic Entrepreneurship Journal*, *2*(4), 317–338.
99. Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic management journal*, *21*(10-11), 1105–1121.
100. Elenkov, D. S. (1997). Strategic Uncertainty and Environmental Scanning: The Case for Institutional Influences On Scanning Behavior. *Strategic management journal*, *18*(4), 287–302.
101. Elhenfnawy, W., Wright, J., Kallepalli, K., Racheal, K., Gupta, A., Parimi, R.,. . . Li, Y. (2016). What Differentiates News Articles with Short and Long Shelf Lives? A Case Study on News Articles at Bloomberg.com. In W. Elhenfnawy, J. Wright, K. Kallepalli, K. Racheal, A. Gupta, R. Parimi,. . . Y. Li (Eds.), *2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom)* (pp. 131–136).
102. Eppink, D. J. (1981). Futures research: Is it used? *Long range planning*, *14*(2), 33–36.
103. Ertek, G., Tapucu, D., & Arın, I. (2013). Text mining with rapidminer. In M. Hofmann (Ed.), *RapidMiner : data mining use cases and business analytics applications* (p. 241). Boca Raton: CRC Press.
104. Everis. (2015). Connected car report. Retrieved from http://www.everis.com/global/WCRepositoryFiles/everis%20connected%20car%20report.pdf
105. Fernhaber, S. A., McDougall, P. P., & Oviatt, B. M. (2007). Exploring the Role of Industry Structure in New Venture Internationalization. *Entrepreneurship Theory and Practice*, *31*(4), 517–542.
106. Fiegenbaum, A., & Karnani, A. (1991). Output flexibility—A competitive advantage for small firms. *Strategic Management Journal*, *12*(2), 101–114.
107. Fisher, G. (2012). Effectuation, Causation, and Bricolage: A Behavioral Comparison of Emerging Theories in Entrepreneurship Research. *Entrepreneurship Theory and Practice*, *36*(5), 1019–1051.
108. Füller, J., Matzler, K., & Hoppe, M. (2008). Brand Community Members as a Source of Innovation. *Journal of Product Innovation Management*, *25*(6), 608–619.
109. Galanakis, K. (2006). Innovation process. Make sense using systems thinking. *Technovation*, *26*(11), 1222–1232.
110. Garcia, R., & Calantone, R. (2002). A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of Product Innovation Management*, *19*(2), 110–132.
111. Geerts, A., Blindenbach-Driessen, F., & Gemmel, P. (2010). Achieving a balance between exploration, exploitation and service firms: A longitudinal study. *Academy of Management Proceedings*, *2010*(1), 1–6.

112. Ghezzi, A., Gastaldi, L., Lettieri, E., Martini, A., & Corso, M. (2016). A role for startups in unleashing the disruptive power of social media. *International Journal of Information Management*, *36*(6), 1152–1159.
113. Gibson, C. B., & Birkinshaw, J. (2004). The antecedents, consequences and mediating role of organizational ambidexterity. *Academy of Management Journal*, *47*(2), 209–226.
114. Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2006). New Venture Growth: A Review and Extension. *Journal of management*, *32*(6), 926–950.
115. Gora, P., & Rüb, I. (2016). Traffic Models for Self-driving Connected Cars. *Transportation Research Procedia*, *14*, 2207–2216.
116. Gordon, R. (2016). ITS and the Connected Vehicle. In R. L. Gordon (Ed.), *Intelligent transportation systems. Functional design for effective traffic management* (pp. 225–235). Cham: Springer.
117. Gordon, T. J. (1992). The Methods of Futures Research. *Annals of the American Academy of Political and Social Science*, *522*, 25–35.
118. Gracht, H. A. von der, Bañuls, V. A., Turoff, M., Skulimowski, A. M., & Gordon, T. J. (2015). Foresight support systems: The future role of ICT for foresight. *Technological Forecasting and Social Change*, *97*, 1–6.
119. Graf, M. (2016). Exploring Sources of Innovation in the Knowledge-Based Economy: The Case of Los Angeles. Retrieved from http://www.rand.org/pubs/rgs_dissertations/RGSD376.html
120. Graham, P. (2013). Age of a startup. Retrieved from http://www.forbes.com/sites/natalierobehmed/2013/12/16/what-is-a-startup/#4ace16764c63
121. Grimpe, C., & Sofka, W. (2009). Search patterns and absorptive capacity: Low- and high-technology sectors in European countries. *Research Policy*, *38*(3), 495–506.
122. Guth, W. D., & Ginsberg, A. (1990). Guest Editors' Introduction: Corporate Entrepreneurship. *Strategic management journal*, *11*, 5–15.
123. Haeckel, S. H. (2004). Peripheral Vision: Sensing and Acting on Weak Signals. *Long range planning*, *37*(2), 181–189.
124. Hambrick, D. C. (1982). Environmental scanning and organizational strategy. *Strategic Management Journal*, *3*(2), 159–174.
125. Hammoud, M. S., & Nash, D. P. (2014). What corporations do with foresight. *European Journal of Futures Research*, *2*(1).
126. Harzing, A.-W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804.
127. Hathaway, I. (2016). What Startup Accelerators Really Do. Retrieved from https://hbr.org/2016/03/what-startup-accelerators-really-do
128. Heger, T., & Rohrbeck, R. (2012). Strategic foresight for collaborative exploration of new business fields. *Technological Forecasting and Social Change*, *79*(5), 819–831.
129. Henkel, J., Rønde, T., & Wagner, M. (2015). And the Winner Is--acquired: Entrepreneurship as a Contest with Acquisition as the Prize.
130. Heuschneider, S., & Herstatt, C. (2016). External search for exploration of future discontinuities and trends: Implications from the literature using co-citation and content analysis. *Technology and Innovation Management, Hamburg University of Technology, Working Paper*. (92).
131. Hidalgo, A., & Albors, J. (2008). Innovation management techniques and tools: a review from theory and practice. *R&D Management*, *38*(2), 113–127.
132. Hill, I. (2015). *Start-up: A Practice Based Guide for New Venture Creation*. London: Palgrave.
133. Hill, S. A., & Birkinshaw, J. (2014). Ambidexterity and Survival in Corporate Venture Units. *Journal of management*, *40*(7), 1899–1931.
134. Hiltunen, E. (2006). Was it a wild card or just our blindness to gradual change. *Journal of Futures Studies*, *11*(2), 61–74.
135. Hiltunen, E. (2008). The future sign and its three dimensions. *Futures*, *40*(3), 247–260.
136. Hippel, E. v. (1988). *The sources of innovation*. New York: Oxford University Press.
137. Hofmann, M., & Chisholm, A. (2016). *Text Mining and Visualization: Case Studies Using Open-Source Tools*. Boca Raton: CRC Press.
138. Hong, J., Shin, J., & Lee, D. (2016). Strategic management of next-generation connected life: Focusing on smart key and car–home connectivity. *Technological Forecasting and Social Change*, *103*, 11–20.
139. Horton, A. (1999). A simple guide to successful foresight. *Foresight*, *1*(1), 5–9.
140. Horváth, A., & Enkel, E. (2014). When general recommendations fail: How to search in single innovation project settings. *R&D Management*, *44*(4), 409–426.
141. Hyundai. (2016). Hyundai Motor Group Announces Roadmap for Connected Car Development - Hyundai Australia. Retrieved from https://www.hyundai.com.au/hyundai-info/news/2016/april/connected-car-development-roadmap

142. Iglesias, J. A., Tiemblo, A., Ledezma, A., & Sanchis, A. (2016). Web news mining in an evolving framework. *Information Fusion*, *28*, 90–98.
143. Ireland, R. D., Hitt, M. A., & Sirmon, D. G. (2003). A Model of Strategic Entrepreneurship: The Construct and its Dimensions. *Journal of management*, *29*(6), 963–989.
144. Isom, C. J., & Jarczyk, D. R. (2009). U.S. Small Business Office of Advocacy: Innovation in Small Business: Drivers of Change and Value - research study. Retrieved from https://www.sba.gov/sites/default/files/rs342tot_0.pdf
145. Jain, S. C. (1984). Environmental scanning in U.S. corporations. *Long range planning*, *17*(2), 117–128.
146. Jenssen, J. I., & NYBAKK, E. (2013). Inter-organizational networks and innovation in small, knowledge-intensive firms: A literature review. *International Journal of Innovation Management*, *17*(02), 1350008/1-1350008/27.
147. Jeong, Y., & Yoon, B. (2015). Development of patent roadmap based on technology roadmap by analyzing patterns of patent development. *Technovation*, *39-40*, 37–52.
148. Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and Problem-Solving Effectiveness in Broadcast Search. *Organization Science*, *21*(5), 1016–1033.
149. Johanning, V., & Mildner, R. (2015). *Car IT kompakt*. Wiesbaden: Springer.
150. Johnson, D. (2001). What is innovation and entrepreneurship?: Lessons for larger organisations. *Industrial and Commercial Training*, *33*(4), 135–140.
151. Kang, K. H., & Kang, J. (2009). How do firms source external knowledge for innovations?: Analysing effects of different knowledge sourcing methods. *International Journal of Innovation Management*, *13*(01), 1–17.
152. Kaserer, M. (2015). Exploring Corporate Foresight in the Start-up Context: Theoretical and Practical Considerations Toward an Integrated Approach.
153. Kayser, V., & Blind, K. (2016). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*.
154. Keller, J., & Gracht, H. A. von der. (2014). ICT and the Foresight Infrastructure of the Future: Effects on the Foresight Discipline. *World Future Review*, *6*(1), 40–47.
155. Kleberger, P. (2015). *On securing the connected car: Methods and protocols for secure vehicle diagnostics*. Zugl.: Göteborg, Univ., Diss., 2015. *Technical report / Department of Computer Science and Engineering, Chalmers University of Technology D: Vol. 120*. Göteborg: Chalmers Univ. of Technology.
156. Kleberger, P., Olovsson, T., & Jonsson, E. (2011). Security aspects of the in-vehicle network in the connected car. In P. Kleberger, T. Olovsson, & E. Jonsson (Eds.), *2011 IEEE Intelligent Vehicles Symposium (IV)* (pp. 528–533).
157. Kohler, T. (2016). Corporate accelerators: Building bridges between corporations and startups. *Business Horizons*, *59*(3), 347–357.
158. Kortum, S., & Lerner, J. (2000). Assessing the contribution of venture capital to innovation. *RAND journal of Economics*, *31*(4), 674–692.
159. Koslowski, T. (2012). Your Connected Vehicle Is Arriving. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/426523/your-connected-vehicle-is-arriving/
160. Kostoff, R. N., Boylan, R., & Simons, G. R. (2004). Disruptive technology roadmaps. *Technological Forecasting and Social Change*, *71*(1-2), 141–159.
161. Kotu, V. (2014). *Predictive analytics and data mining: Concepts and practice with rapidminer* (1st edition). Waltham: Elsevier.
162. Kuratko, D. F., & Audretsch, D. B. (2013). Clarifying the domains of corporate entrepreneurship. *International Entrepreneurship and Management Journal*, *9*(3), 323–335.
163. Kuratko, D. F., Hornsby, J. S., & Hayton, J. (2015). Corporate entrepreneurship: The innovative challenge for a new global economic reality. *Small Business Economics*, *45*(2), 245–253.
164. Kuusi, O., Cuhls, K., & Steinmüller, K. (2015). The futures Map and its quality criteria. *European Journal of Futures Research*, *3*(22), 1–14.
165. Laursen, K., Leone, M. I., & Torrisi, S. (2010). Technological exploration through licensing: New insights from the licensee's point of view. *Industrial and Corporate Change*, *19*(3), 871–897.
166. Laursen, K., & Salter, A. (2004). Searching high and low: What types of firms use universities as a source of innovation? *Research Policy*, *33*(8), 1201–1215.
167. Lee, C., Kang, B., & Shin, J. (2015). Novelty-focused patent mapping for technology opportunity analysis. *Technological Forecasting and Social Change*, *90*, 355–365.
168. Lee, G. M., Crespi, N., Choi, J. K., & Boussard, M. (2013). Internet of Things. In E. Bertin, N. Crespi, & T. Magedanz (Eds.), *Evolution of telecommunication services. The convergence of telecom and internet: technologies and ecosystems* (pp. 257–282). Berlin: Springer.
169. Leiponen, A., & Helfat, C. E. (2009). Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic Management Journal*, *31*, 224-236.

170. Liang, Y. E., & Yuan, S.-T. D. (2016). Predicting investor funding behavior using crunchbase social network features. *Internet Research*, *26*(1), 74–100.

171. Lim, K., Chesbrough, H., & Ruan, Y. (2010). Open innovation and patterns of R&D competition. *International Journal of Technology Management*, *52*(3/4), 295.

172. Lu, N., Cheng, N., Zhang, N., Shen, X., & Mark, J. W. (2014). Connected Vehicles: Solutions and Challenges. *IEEE Internet of Things Journal*, *1*(4), 289–299.

173. Lucas, H. C., & Goh, J. M. (2009). Disruptive technology: How Kodak missed the digital photography revolution. *The Journal of Strategic Information Systems*, *18*(1), 46–55.

174. Luger, M. I., & Koo, J. (2005). Defining and Tracking Business Start-Ups. *Small Business Economics*, *24*(1), 17–28.

175. Maine, E., Soh, P.-H., & Dos Santos, N. (2015). The role of entrepreneurial decision-making in opportunity creation and recognition. *Technovation*, *39-40*, 53–72.

176. March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

177. Markides, C. (2006). Disruptive innovation: In need of better theory. *Journal of Product Innovation Management*, *23*(1), 19–25.

178. Martin, B. R. (1995). Foresight in science and technology. *Technology Analysis & Strategic Management*, *7*(2), 139–168.

179. Martín-de Castro, G., & Fischer, A. (2011). Recognizing opportunities: Initiating service innovation in PSFs. *Journal of Knowledge Management*, *15*(6), 915–927.

180. Mason, R. O. (1969). A dialectical approach to strategic planning. *Management Science*, *15*(8), 403–414.

181. McGrath, R. G., & Ming-Hone Tsai. (1996). Innovation, Competitive Advantage and Rent: A Model and Test. *Management Science*, *42*(3), 389–403.

182. McMaster, M. (1996). Foresight: Exploring the structure of the future. *Long range planning*, *29*(2), 149–155.

183. Mendonça, S., Pina e Cunha, M., Kaivo-oja, J., & Ruff, F. (2004). Wild cards, weak signals and organisational improvisation. *Futures*, *36*(2), 201–218.

184. Mian, S., Lamine, W., & Fayolle, A. (2016). Technology Business Incubation: An overview of the state of knowledge. *Technovation*, *50-51*, 1–12.

185. Minelli, S., Izadpanah, P., & Razavi, S. (2015). Evaluation of connected vehicle impact on mobility and mode choice. *Journal of Traffic and Transportation Engineering (English Edition)*, *2*(5), 301–312.

186. Mintzberg, H. (1994). The fall and rise of strategic planning. *Harvard Business Review*, *72*(1), 107.

187. Miranda Santo, M. de, Coelho, G. M., dos Santos, D. M., & Filho, L. F. (2006). Text mining as a valuable tool in foresight exercises: A study on nanotechnology. *Technological Forecasting and Social Change*, *73*(8), 1013–1027.

188. Müller, A. W. (2008). *Strategic Foresight-Prozesse strategischer Trend-und Zukunftsforschung in Unternehmen.*

189. Narayanan, V. K., Yang, Y., & Zahra, S. A. (2009). Corporate venturing and value creation: A review and proposed framework. *Research Policy*, *38*(1), 58–76.

190. Nikander, I. O., & Eloranta, E. (1997). Preliminary signals and early warnings in industrial investment projects. *International Journal of Project Management*, *15*(6), 371–376.

191. Norman, D. A., & Verganti, R. (2014). Incremental and radical innovation: Design research vs. technology and meaning change. *Design Issues*, *30*(1), 78–96.

192. O'Connor, G. C., & Rice, M. P. (2001). Opportunity Recognition and Breakthrough Innovation in Large Established Firms. *California Management Review*, *43*(2), 95–116.

193. Oecd. (2005). *Oslo Manual*. Paris: OECD Publishing:.

194. Onan, A., Korukoğlu, S., & Bulut, H. (2016). Ensemble of keyword extraction methods and classifiers in text classification. *Expert Systems with Applications*, *57*, 232–247.

195. O'Reilly, C. A., & Tushman, M. L. (2013). Organizational Ambidexterity: Past, Present, and Future. *Academy of Management Perspectives*, *27*(4), 324–338.

196. Ortt, J. R., & van der Duin, Patrick A. (2008). The evolution of innovation management towards contextual innovation. *European Journal of Innovation Management*, *11*(4), 522–538.

197. Oviatt, B. M., & McDougall, P. P. (1994). Toward a Theory of International New ventures. *Journal of International Business Studies*, *25*(1), 45–64.

198. Paliokaitė, A., & Pačėsa, N. (2015). The relationship between organisational foresight and organisational ambidexterity. *Technological Forecasting and Social Change*, *101*, 165–181.

199. Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, *52*(2), 183–199.

200. Pauwels, C., Clarysse, B., Wright, M., & van Hove, J. (2016). Understanding a new generation incubation model: The accelerator. *Technovation*, *50-51*, 13–24.

201. Peng, T., Liu, L., & Zuo, W. (2014). PU text classification enhanced by term frequency-inverse document frequency-improved weighting. *Concurrency and Computation: Practice and Experience*, *26*(3), 728–741.

202. Petersen, J. L. (1997). The "wild cards" in our future: preparing for the improbable. *The Futurist*, *31*(4), 43.

203. Phan, P. H., Wright, M., Ucbasaran, D., & Tan, W.-L. (2009). Corporate entrepreneurship: Current research and future directions. *Journal of Business Venturing*, *24*(3), 197–205.

204. Phelps, C., Heidl, R., & Wadhwa, A. (2012). Knowledge, Networks, and Knowledge Networks: A Review and Research Agenda. *Journal of management*, *38*(4), 1115–1166.

205. Piirainen, K. A., Gonzalez, R. A., & Bragge, J. (2012). A systemic evaluation framework for futures research. *Futures*, *44*(5), 464–474.

206. Pina e Cunha, M., & Chia, R. (2007). Using Teams to Avoid Peripheral Blindness. *Long range planning*, *40*(6), 559–573.

207. Pittaway, L., Robertson, M., Munir, K., Denyer, D., & Neely, A. (2004). Networking and innovation: A systematic review of the evidence. *International Journal of Management Reviews*, *5-6*(3-4), 137–168.

208. Plehn-Dujowich, J. M. (2013). Product Innovations by Young and Small Firms research study. Retrieved from https://www.sba.gov/sites/default/files/files/rs408tot.pdf

209. Porter, M. E. (1985). Technology and Competitive Advantage. *Journal of Business Strategy*, *5*(3), 60–78.

210. Porter, M. E. (1990). The competitive advantage of nations. *Harvard Business Review*, *68*(2), 73.

211. Porter, M. E. (1996). What is strategy?(management strategy). *Harvard Business Review*, *74*(6), 61.

212. Prahalad, C. K. (2004). The Blinders of Dominant Logic. *Long range planning*, *37*(2), 171–179.

213. Raincomplex. (2013). Euclidean Voronoi Diagram. Retrieved from https://commons.wikimedia.org/wiki/File:Euclidean_Voronoi_Diagram.png#mediaviewer/%20File:Euclidean_Voronoi_Diagram.png

214. Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance. *Organization Science*, *20*(4), 685–695.

215. RapidMiner. (2016). RapidMiner Studio 7.3. Retrieved from https://rapidminer.com/

216. Rayes, A., & Salam, S. (2017). Internet of Things (IoT) Overview. In A. Rayes & S. Salam (Eds.), *Internet of Things -- from hype to reality. The road to digitization* (pp. 1–34). Cham: Springer.

217. Read, S., Song, M., & Smit, W. (2009). A meta-analytic review of effectuation and venture performance. *Journal of Business Venturing*, *24*(6), 573–587.

218. Reid, S. E., Roberts, D., & Moore, K. (2015). Technology Vision for Radical Innovation and Its Impact on Early Success. *Journal of Product Innovation Management*, *32*(4), 593–609.

219. Remane, G., Hildebrandt, B., Hanelt, A., & Kolbe, L. M. (Eds.) 2016. *Discovering New Digital Business Model Types - A Study of Technology Startups from the Mobility Sector.*

220. Reymen, I. M., Andries, P., Berends, H., Mauer, R., Stephan, U., & van Burg, E. (2015). Understanding Dynamics of Strategic Decision Making in Venture Creation: A Process Study of Effectuation and Causation. *Strategic Entrepreneurship Journal*, *9*(4), 351–379.

221. Richter, M. (2013). Business model innovation for sustainable energy: German utilities and renewable energy. *Energy Policy*, *62*, 1226–1237.

222. Ripsas, S., & Tröger, S. (2014). Deutscher Startup Monitor. Retrieved from http://deutscherstartupmonitor.de/fileadmin/dsm/dsm-14/DSM_2014.pdf

223. Ritter, T., & Gemünden, H. G. (2003). Network competence. *Journal of Business Research*, *56*(9), 745–755.

224. Robertson, T. S. (1967). The Process of Innovation and the Diffusion of Innovation. *Journal of Marketing*, *31*(1), 14.

225. Rockfellow, J. D. (1994). Wild cards: preparing for "the big one". *The Futurist*, *28*(1), 14.

226. Roessler, M., & Velamuri, V. K. (Eds.) 2015. *Corporate Incubation as a Tool to Foster Business Model Innovation*: The International Society for Professional Innovation Management (ISPIM) Conference.

227. Rohrbeck, R. (2010a). *Corporate Foresight: Towards a Maturity Model for the Future Orientation of a Firm*. Heidelberg: Physica-Verlag HD.

228. Rohrbeck, R. (2010b). Harnessing a network of experts for competitive advantage: Technology scouting in the ICT industry. *R&D Management*, *40*(2), 169–180.

229. Rohrbeck, R., & Bade, M. (2012). Environmental scanning, futures research, strategic foresight and organizational future orientation: a review, integration, and future research directions. In R. Rohrbeck & M. Bade (Eds.), *Environmental scanning, futures research, strategic foresight and organizational future orientation: a review, integration, and future research directions.*

230. Rohrbeck, R., Battistella, C., & Huizingh, E. (2015). Corporate foresight: an emerging field with a rich tradition. *Technological Forecasting and Social Change*, *101*, 1–9.

231. Rohrbeck, R., & Gemünden, H. G. (2011). Corporate foresight: Its three roles in enhancing the innovation capacity of a firm. *Technological Forecasting and Social Change*, *78*(2), 231–243.
232. Rohrbeck, R., & Schwarz, J. O. (2013). The value contribution of strategic foresight: Insights from an empirical study of large European companies. *Technological Forecasting and Social Change*, *80*(8), 1593–1606.
233. Rohrbeck, R., Thom, N., & Arnold, H. (2015). IT tools for foresight: The integrated insight and response system of Deutsche Telekom Innovation Laboratories. *Technological Forecasting and Social Change*, *97*, 115-126.
234. Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic Keyword Extraction from Individual Documents. In J. Kogan & M. W. Berry (Eds.), *Text mining. Applications and theory* (pp. 1–20). Chichester: Wiley.
235. Rothwell, R. (1984). The role of small firms in the emergence of new technologies. *Omega*, *12*(1), 19–29.
236. Rowe, F. (2014). What literature review is not: Diversity, boundaries and recommendations. *European Journal of Information Systems*, *23*(3), 241–255.
237. Ruff, F. (2015). The advanced role of corporate foresight in innovation and strategic management—Reflections on practical experiences from the automotive industry. *Technological Forecasting and Social Change*, *101*, 37–48.
238. Sangster, A. (2015). You Cannot Judge a Book by Its Cover: The Problems with Journal Rankings. *Accounting Education*, *24*(3), 175–186.
239. Sarasvathy, S. D. (2001). Causation and Effectuation: Toward a Theoretical Shift From Economic Inevitability to Entrepreneurial Contingency. *Academy of Management Review*, *26*(2), 243–263.
240. Sarasvathy, S., Kumar, K., York, J. G., & Bhagavatula, S. (2014). An Effectual Approach to International Entrepreneurship: Overlaps, Challenges, and Provocative Possibilities. *Entrepreneurship Theory and Practice*, *38*(1), 71–93.
241. Sardar, Z. (2010). The Namesake: Futures; futures studies; futurology; futuristic; foresight—What's in a name? *Futures*, *42*(3), 177–184.
242. Satapathy, S. C., Bhateja, V., & Joshi, A. (Eds.). (2017). *Advances in Intelligent Systems and Computing*. *Proceedings of the International Conference on Data Engineering and Communication Technology*. Singapore: Springer.
243. Saunders, M., Lewis, P., & Thornhill, A. (2011). *Research methods for business students*. Harlow: Pearson Education.
244. Schiele, H. (2010). Early supplier integration: the dual role of purchasing in new product development. *R&D Management*, *40*(2), 138–153.
245. Schildt, H. A., Zahra, S. A., & Sillanpaa, A. (2006). Scholarly Communities in Entrepreneurship Research: A Co-Citation Analysis. *Entrepreneurship Theory and Practice*, *30*(3), 399–415.
246. Schryen, G. (2013). Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, *22*(2), 139–169.
247. Schulze, A., MacDuffie, J. P., & Täube, F. A. (2015). Introduction: knowledge generation and innovation diffusion in the global automotive industry—change and stability during turbulent times. *Industrial and Corporate Change*, *24*(3), 603–611.
248. Schumpeter, J. A. (1912). *Theorie der wirtschaftlichen Entwicklung*. Leipzig: Duncker & Humblot.
249. Schumpeter, J. A. (2004). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. New Brunswick: Transaction publishers (Original work published 1934).
250. Scozzi, B., Garavelli, C., & Crowston, K. (2005). Methods for modeling and supporting innovation processes in SMEs. *European Journal of Innovation Management*, *8*(1), 120–137.
251. Seo, W., Yoon, J., Park, H., Coh, B.-y., Lee, J.-M., & Kwon, O.-J. (2016). Product opportunity identification based on internal capabilities using text mining and association rule mining. *Technological Forecasting and Social Change*, *105*, 94–104.
252. Sharma, P., & Chrisman, J. J. (1999). Toward a Reconciliation of the Definitional Issues in the Field of Corporate Entrepreneurship. *Entrepreneurship: Theory and Practice*, *23*(3), 11.
253. Shraddha, K., & Emmanuel, M. (2014). Review and Comparative Study of Clustering Techniques. *International Journal of Computer Science and Information Technology*, *5*(1), 805–812.
254. Siddiqi, S. (2015). Keyword and Keyphrase Extraction Techniques: A Literature Review. *International Journal of Computer Applications*, *109*(2), 975–8887.
255. Siegel, R., Siegel, E., & MacMillan, I. C. (1988). Corporate venture capitalists: Autonomy, obstacles, and performance. *Journal of Business Venturing*, *3*(3), 233–247.
256. Sinclair, S., & Rockwell, G. (2016). Voyant Tools. Retrieved from https://voyant-tools.org/docs/#!/guide/about
257. Sirmon, D. G., Hitt, M. A., Arregle, J.-L., & Campbell, J. T. (2010). The dynamic interplay of capability strengths and weaknesses: Investigating the bases of temporary competitive advantage. *Strategic Management Journal*, *31*(13), 1386–1409.

258.Slama, D., Puhlmann, F., Morrish, J., & Bhatnagar, R. M. (2015). *Enterprise IoT: Strategies and Best Practices for Connected Products and Services* (ss). Beijing: O'Reilly.

259.Slaughter, R. A. (1990). Assessing the quest for future knowledge. *Futures*, *22*(2), 153–166.

260.Slaughter, R. A. (1996). Futures studies. *Futures*, *28*(8), 751–762.

261.Sood, A., & Tellis, G. J. (2011). Demystifying Disruption: A New Model for Understanding and Predicting Disruptive Technologies. *Marketing Science*, *30*(2), 339–354.

262.Sorescu, A. B., Chandy, R. K., & Prabhu, J. C. (2003). Sources and financial consequences of radical innovation: insights from pharmaceuticals. *Journal of Marketing*, *67*(4), 82.

263.Spithoven, A., Clarysse, B., & Knockaert, M. (2010). Building absorptive capacity to organise inbound open innovation in traditional industries. *Technovation*, *30*(2), 130–141.

264.SpringerProfessional. (2016). Fachwissen. Digital. Finden. Retrieved from https://www.springerprofessional.de/

265.Sreenivasulu, G., Raju, S. V., & Rao, N. S. (2017). Review of Clustering Techniques. In S. C. Satapathy, V. Bhateja, & A. Joshi (Eds.), *Advances in Intelligent Systems and Computing. Proceedings of the International Conference on Data Engineering and Communication Technology* (pp. 523–535). Singapore: Springer.

266.Stettner, U., & Lavie, D. (2014). Ambidexterity under scrutiny: Exploration and exploitation via internal organization, alliances, and acquisitions. *Strategic Management Journal*, *35*(13), 1903–1929.

267.Stevenson, H. H., & Jarillo, J. C. (1990). A Paradigm of Entrepreneurship: Entrepreneurial Management. *Strategic Management Journal*, *11*, 17–27.

268.Swan, M. (2015). Connected Car: Quantified Self becomes Quantified Car. *Journal of Sensor and Actuator Networks*, *4*(1), 2–29.

269.Talay, M. B., & Townsend, J. D. (2015). Do or die: competitive effects and Red Queen dynamics in the product survival race. *Industrial and Corporate Change*, *24*(3), 721–738.

270.Taleb, N. N. (2010). *The black swan: The impact of the highly improbable* (2., rev. ed.). London: Penguin.

271.Tang, L., & Koveos, P. E. (2004). Venture entrepreneurship, innovation entrepreneurship, and economic growth. *Journal of Developmental Entrepreneurship*, *9*(2), 161.

272.Technology Futures Analysis Methods Working Group. (2004). Technology futures analysis: Toward integration of the field and new methods. *Technological Forecasting and Social Change*, *71*(3), 287–303.

273.Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic management journal*, *28*(13), 1319–1350.

274.Teece, D. J. (2014). A dynamic capabilities-based entrepreneurial theory of the multinational enterprise. *Journal of International Business Studies*, *45*(1), 8–37.

275.Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, *18*(7), 509–533.

276.Tellis, G. J., Prabhu, J. C., & Chandy, R. C. (2009). Radical innovation across nations: the preeminence of corporate culture. *Journal of Marketing*, *73*(1), 3.

277.Terwiesch, C., & Xu, Y. (2008). Innovation Contests, Open Innovation, and Multiagent Problem Solving. *Management Science*, *54*(9), 1529–1543.

278.Thorleuchter, D., & van den Poel, D. (2015). Idea mining for web-based weak signal detection. *Futures*, *66*, 25–34.

279.Thorleuchter, D., van den Poel, D., & Prinzie, A. (2010). Mining ideas from textual information. *Expert Systems with Applications*, *37*(10), 7182–7188.

280.Tidd, J. (2001). Innovation management in context: environment, organization and performance. *International Journal of Management Reviews*, *3*(3), 169–183.

281.Tödtling, F., Lehner, P., & Kaufmann, A. (2009). Do different types of innovation rely on specific kinds of knowledge interactions? *Technovation*, *29*(1), 59–71.

282.Townsend, J. D., & Calantone, R. J. (2014). Evolution and transformation of innovation in the global automotive industry. *Journal of Product Innovation Management*, *31*(1), 4–7.

283.Tran, T. N., Wehrens, R., & Buydens, L. M. (2006). KNN-kernel density-based clustering for high-dimensional multivariate data. *Computational Statistics & Data Analysis*, *51*(2), 513–525.

284.Tranfield, D., Denyer, D., Marcos, J., & Burr, M. (2004). Co-producing management knowledge. *Management Decision*, *42*(3/4), 375–386.

285.Tushman, M., Smith, W. K., Wood, R. C., Westerman, G., & O'Reilly, C. (2010). Organizational designs and innovation streams. *Industrial and Corporate Change*, *19*(5), 1331–1366.

286.Tushman, M. L., & O'Reilly, Charles A., III. (1996). Ambidextrous organizations: managing evolutionary and revolutionary change. *California Management Review*, *38*(4), 8.

287.Tyssen, M. (2012). *Zukunftsorientierung und dynamische Fähigkeiten*. Wiesbaden: Gabler Verlag.

288. Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing*, *26*(3), 341–358.
289. University of Twente Library. (2013). ITC - Professional journals. Retrieved from https://www.itc.nl/Pub/Home/library/Library-Guides/LiteracyCourse/Module_2/Module2-Professional_journals.html
290. Urban, B., & Wood, E. (2015). The importance of opportunity recognition behaviour and motivators of employees when engaged in corporate entrepreneurship. *Journal of Business Economics and Management*, *16*(5), 980–994.
291. van Burg, E., Podoynitsyna, K., Beck, L., & Lommelen, T. (2012). Directive Deficiencies: How Resource Constraints Direct Opportunity Identification in SMEs. *Journal of Product Innovation Management*, *29*(6), 1000–1011.
292. van der Duin, P. (2006). *Qualitative futures research for innovation*. Delft: Eburon.
293. van der Duin, P. (2014). The Crystal Ball is not a Black Box: Futures Research in Scientific and Organizational Perspective. *Journal of Futures Studies*, *19*(2), 125–134.
294. van der Duin, P., Heger, T., & Schlesinger, M. D. (2014). Toward networked foresight?: Exploring the use of futures research in innovation networks. *Futures*, *59*, 62–78.
295. van Praag, C. M., & Versloot, P. H. (2007). What is the value of entrepreneurship? A review of recent research. *Small business economics*, *29*(4), 351–382.
296. Varis, M., & Littunen, H. (2010). Types of innovation, sources of information and performance in entrepreneurial SMEs. *European Journal of Innovation Management*, *13*(2), 128–154.
297. Vecchiato, R. (2015). Creating value through foresight: First mover advantages and strategic agility. *Technological Forecasting and Social Change*, *101*, 25–36.
298. Vecchiato, R., & Roveda, C. (2010). Strategic foresight in corporate organizations: Handling the effect and response uncertainty of technology and social drivers of change. *Technological Forecasting and Social Change*, *77*(9), 1527–1539.
299. Venkatraman, N., Lee, C.-H., & Iyer, B. (2007). Strategic ambidexterity and sales growth: A longitudinal test in the software sector. *N.d.*, pp. 1–45.
300. Veryzer, R. W. (1998). Discontinuous Innovation and the New Product Development Process. *Journal of Product Innovation Management*, *15*(4), 304–321.
301. Volkswagen AG. (2016). Volkswagen announces strategic partnership with Gett. Retrieved from http://media.vw.com/release/1197/
302. Vossen, R. W. (1998). Relative strengths and weaknesses of small firms in innovation. *International Small Business Journal*, *16*(3), 88.
303. Wack, P. (1985). Scenarios: uncharted waters ahead. *Harvard Business Review*, *63*(5), 73–89.
304. Wallin, M. W., & Krogh, G. von. (2010). Organizing for Open Innovation, *39*(2), 145–154.
305. Wan, F., Williamson, P. J., & Yin, E. (2015). Antecedents and implications of disruptive innovation: Evidence from China. *Technovation*, *39*, 94–104.
306. Wang, C. L., Senaratne, C., & Rafiq, M. (2015). Success Traps, Dynamic Capabilities and Firm Performance. *British Journal of Management*, *26*(1), 26–44.
307. Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly*, *26*(2), xiii.
308. Weiblen, T., & Chesbrough, H. W. (2015). Engaging with Startups to Enhance Corporate Innovation. *California Management Review*, *57*(2), 66–90.
309. Werhahn, D., Mauer, R., Flatten, T. C., & Brettel, M. (2015). Validating effectual orientation as strategic direction in the corporate context. *European Management Journal*, *33*(5), 305–313.
310. West, J., & Bogers, M. (2014). Leveraging External Sources of Innovation: A Review of Research on Open Innovation. *Journal of Product Innovation Management*, *31*(4), 814–831.
311. West, J., & Gallagher, S. (2006). Challenges of open innovation: the paradox of firm investment in open-source software. *R&D Management*, *36*(3), 319–331.
312. West, J., Salter, A., Vanhaverbeke, W., & Chesbrough, H. (2014). Open innovation: The next decade. *Research Policy*, *43*(5), 805–811.
313. Willett, P. (2006). The Porter stemming algorithm: Then and now. *Program*, *40*(3), 219–223.
314. Winter, S. G. (2004). Specialised Perception, Selection, and Strategic Surprise: Learning from the Moths and Bees. *Long range planning*, *37*(2), 163–169.
315. Wissema, J. G. (1981). Futures research—is it useful? *Long range planning*, *14*(2), 29–32.
316. Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45–55.
317. Yoo, J.-Y., & Yang, D. (2015). Classification Scheme of Unstructured Text Document using TF-IDF and Naive Bayes Classifier. In J.-Y. Yoo & D. Yang (Eds.): *Advanced Science and Technology Letters, Computer and Computing Science 2015* (pp. 263–266). Science & Engineering Research Support soCiety.

318. Yoon, B., & Park, Y. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. *The Journal of High Technology Management Research*, *15*(1), 37–50.

319. Yoon, B., Park, I., & Coh, B.-y. (2014). Exploring technological opportunities by linking technology and products: Application of morphology analysis and text mining. *Technological Forecasting and Social Change*, *86*, 287–303.

320. Yoon, B., Phaal, R., & Probert, D. (2008). Morphology analysis for technology roadmapping: Application of text mining. *R&D Management*, *38*(1), 51–68.

321. Yoon, J. (2012). Detecting weak signals for long-term business opportunities using text mining of Web news. *Expert Systems with Applications*, *39*(16), 12543–12550.

322. Yoon, J., Park, H., & Kim, K. (2013). Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis. *Scientometrics*, *94*(1), 313–331.

323. Yoon, J., Park, H., Seo, W., Lee, J.-M., Coh, B.-y., & Kim, J. (2015). Technology opportunity discovery (TOD) from existing technologies and products: A function-based TOD framework. *Technological Forecasting and Social Change*, *100*, 153–167.

324. Yoon, J., Seo, W., Coh, B.-y., Song, I., & Lee, J.-M. (2016). Identifying product opportunities using collaborative filtering-based patent analysis. *Computers & Industrial Engineering*.

325. Yu, D., & Hang, C. C. (2010). A reflective review of disruptive innovation theory. *International Journal of Management Reviews*, *12*(4), 435–452.

326. Zahra, S., Sapienza, H. J., & Davidsson, P. (2006). Entrepreneurship and Dynamic Capabilities: A Review, Model and Research Agenda. *Journal of Management Studies*, *43*(4), 917–955.

327. Zahra, S. A. (2015). Corporate entrepreneurship as knowledge creation and conversion: The role of entrepreneurial hubs. *Small Business Economics*, *44*(4), 727–735.

328. Zahra, S. A., & George, G. (2002). Absorptive capacity: a review, reconceptualization, and extension. *Academy of Management Review*, *27*(2), 185.

329. Zahra, S. A., Korri, J. S., & Yu, J. (2005). Cognition and international entrepreneurship: Implications for research on international opportunity recognition and exploitation. *International Business Review*, *14*(2), 129–146.

330. Zhang, Y., Zhang, G., Chen, H., Porter, A. L., Zhu, D., & Lu, J. (2016). Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. *Technological Forecasting and Social Change*, *105*, 179–191.

331. Zhao, F. (2005). Exploring the synergy between entrepreneurship and innovation. *International Journal of Entrepreneurial Behavior & Research*, *11*(1), 25–41.

332. Zhou, H., Sandner, P. G., Martinelli, S. L., & Block, J. H. (2016). Patents, trademarks, and their complementarity in venture capital funding. *Technovation*, *47*, 14–22.

333. Zhou, K. Z., & Wu, F. (2009). Technological capability, strategic flexibility, and product innovation. *Strategic Management Journal*, *31*(5), 547–561.

Appendices

Appendix 1 - Classification of innovation

Reference	Classification	Description of classification
(Robertson, 1967, pp. 15–16)	Continuous	*"A continuous innovation has the least disrupting influence on established patterns. Alteration of a product is involved, rather than the establishment of a new product. Examples: fluoride toothpaste; new-model automobile changeovers; menthol cigarettes."*
	Dynamically continuous	*"A dynamically continuous innovation has more disrupting effects than a continuous innovation, although it still does not generally alter established patterns. It may involve the creation of a new product or the alteration of an existing product. Examples: electric toothbrushes; the Mustang automobile; Touch-Tone telephones."*
	Discontinuous	*"A discontinuous innovation involves the establishment of a new product and the establishment of new behavior patterns. Examples: television; computers."*
(Porter, 1985, p. 77)	Continuous, incremental evolution	*"Where there is incremental technological change, the process is more likely to be determined by actions of industry participants or spin-offs of these participants."*
	Discontinuous evolution	*"Where there is technological discontinuity, the sources of technology are much more likely to be outside the industry."*
(Abernathy & Clark, 1985, pp. 7–12)	Conservative	*"On the conservative end of the scale are those innovations that serve to enhance the value or applicability of the firm's existing competence. Clearly, all technological innovation imposes change of some kind, but change need not be destructive. Innovation in product technology may solve problems or eliminate flaws in a design that makes existing channels of distribution more a tractive and effective."*
	Radical	*"On the radical end of the scale, the effect of innovation is quite the opposite. Instead of enhancing and strengthening, innovation of this sort disrupts and destroys. It changes the technology*

		of process or product in a way that imposes requirements that the existing resources, skills and knowledge satisfy poorly or not at all."
(Dewar & Dutton, 1986, pp. 1422–1423)	Radical innovation	*"Radical innovations are fundamental changes that represent revolutionary change in technology. They represent clear departures from existing practice."*
	Incremental innovation	*"[…] incremental innovations are minor improvements or simple adjustments in current technology."*
(Veryzer, 1998, p. 307)	Continuous	*"[…] encompasses products that utilize existing technology and provide the same benefits as existing products. Such products are continuous both in terms of the technology employed and the way they are experienced by customers. Although they may be new they are not very innovative."*
	Technologically discontinuous	*"In addition to continuous new products, new products may be discontinuous with respect to technology, the benefits perceived by the customer, or both."*
	Commercially discontinuous	*"Products that are perceived by customers as being really new regardless of whether or not they utilize new technology are commercially discontinuous."*
	Technologically and commercially discontinuous	*"In addition to continuous new products, new products may be discontinuous with respect to technology, the benefits perceived by the customer, or both. In cases where the delivery of new benefits involves the application of a significant new technology, the product is technologically discontinuous in addition to being commercially discontinuous."*
(Chandy & Tellis, 1998, p. 476)	Incremental innovation	*"Incremental innovations involve relatively minor changes in technology and provide relatively low incremental customer benefits per dollar."*
	Market breakthroughs	*"Market breakthroughs are based on core technology that is similar to existing products but provide substantially higher customer benefits per dollar."*

	Technological breakthroughs	*"Technological breakthroughs adopt a substantially different technology than existing products but do not provide superior customer benefits per dollar."*
	Radical innovation	*"[…] radical innovations involve substantially new technology and provide substantially greater customer benefits per dollar, relative to existing products."*
(Assink, 2006, pp. 217–218)	Incremental or sustainable innovation (remodeling functionality)	*"Incremental innovation development remains within the boundaries of the existing market and technology or processes of an organisation and carries lower financial and market-acceptance risks."*
	Radical or disruptive innovations	*"A successfully exploited radical new product, process, or concept that significantly transforms the demand and needs of an existing market or industry, disrupts its former key players and creates whole new business practices or markets with significant societal impact."*
	Breakthrough innovations	*"Breakthrough innovations are based on inventions that serve as a source of many subsequent inventions."*
(Corso & Pellegrini, 2007, 376-342)	Incremental exploitation	*"Incremental exploitation refers to the 'exploitation' capabilities that lead to incremental product and process improvement. It deals with continuous improvement (CI), or kaizen […]"*
	Incremental exploration	*"Incremental exploration refers to exploration capabilities that lead to incremental innovation – i.e., innovation which builds on existing knowledge – in both processes and products.* *As far as processes are concerned, it is important to mention business process re-design or re-engineering (BPR) that attempts to contribute to performance improvement, but, in contrast to CI, through non-linear changes and without the widespread involvement of employees."*
	Radical exploration	*"Radical exploration refers to exploration capabilities that lead to radical innovation in new and uncharted territories. This stream of the literature deals with discontinuities, characterized by high instability (unpredictable and unstable conditions) and high uncertainty (the extent to which knowledge can be acquired to help deal with the environment)."*

	Radical exploitation	*"Radical exploitation refers to exploitation capabilities that lead to radical innovation in new and uncharted territories. Within this stream of the literature, many contributions highlight that in different industries many incumbents, while leveraging on their existing knowledge, could be credited with many radical innovations."*
(Norman & Verganti, 2014, pp. 82–84)	Incremental innovation	*"Incremental product innovation refers to the <u>small changes</u> in a product that help to <u>improve its performance, lower its costs, and enhance its desirability, or simply result in a new model release</u>."*
	Radical innovation	*"It is often characterized as disruptive or <u>competence destroying, or as breakthrough</u>, with all these labels sharing the same concept that radical innovation implies a <u>discontinuity with the past</u>."*
(Christensen, 1997/2011, xviii; Christensen & Raynor, 2013, pp. 39–40)	Sustaining technologies/innovation	*"Some sustaining technologies can be discontinuous or radical in character, while others are of an incremental nature. What all sustaining technologies have in common is that they improve the performance that mainstream customers in major markets have historically valued."*
	Disruptive technologies/innovation	*"Disruptive technologies bring to a market a very different value proposition than had been available previously. Generally, disruptive technologies underperform established products in mainstream markets. But they have other feature that a few fringe (and generally new) customers value. Products based on disruptive technologies are typically cheaper, simpler, smaller, and, frequently, more convenient to use."*

Appendix 2 - TF-IDF value matrix (RapidMiner calculation)

ExampleSet (238 examples, 4 special attributes, 4007 regular attributes) Filter (238 / 238 examples): all

Row No.	label	metadata_file	metadata_d...	metadata_p...	abil	abl	abnorm	abroad	absolut	abstract	academi	acceler
1	Text Mining P...	6361335687...	Oct 29, 2016 ...	C:\Users\PM\...	0	0.012	0	0	0	0	0	0
2	Text Mining P...	6361335834...	Oct 29, 2016 ...	C:\Users\PM\...	0.025	0.015	0	0	0	0	0	0
3	Text Mining P...	6361394465...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.016	0	0	0	0	0	0.060
4	Text Mining P...	6361394740...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.015	0	0	0	0	0	0
5	Text Mining P...	6361394785...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.005	0	0	0	0	0	0
6	Text Mining P...	6361394789...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.006	0	0	0.103	0	0	0.045
7	Text Mining P...	6361394824...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.004	0	0	0	0	0	0
8	Text Mining P...	6361394833...	Nov 5, 2016 1...	C:\Users\PM\...	0.014	0.004	0	0	0	0	0	0.031
9	Text Mining P...	6361394844...	Nov 5, 2016 1...	C:\Users\PM\...	0	0.006	0	0	0	0	0	0.022
10	Text Mining P...	Almeida, Serr...	Oct 29, 2016 ...	C:\Users\PM\...	0	0.006	0	0	0	0.058	0	0
11	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.002	0.026	0	0	0	0	0
12	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.015	0	0	0	0	0	0	0
13	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.003	0.003	0	0	0	0	0	0.012
14	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.004	0.003	0	0	0	0	0	0
15	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.002	0	0	0	0	0	0
16	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.003	0	0	0	0	0	0	0
17	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.006	0	0	0	0	0	0	0
18	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.002	0	0	0	0	0	0
19	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.003	0	0	0	0	0	0
20	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.010	0	0	0	0	0	0
21	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0	0	0	0	0	0	0
22	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0	0	0	0	0	0	0
23	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0.017	0.010	0	0	0	0	0.005
24	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0	0	0	0	0	0	0.003
25	Text Mining P...	art 10.1007_...	Nov 5, 2016 2...	C:\Users\PM\...	0	0	0	0	0	0	0	0
26	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0.020	0.005	0	0	0	0	0	0
27	Text Mining P...	art 10.1007_...	Nov 6, 2016 1...	C:\Users\PM\...	0	0	0	0	0	0	0	0

Appendix 3 - Similarity matrix

	A	H	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
1	**Startup name**	**Startup description**	**Similarity index**	ac	ad	ad	al	an	an	at	at	au	au	au	au	ba	ba	be	be	bu	bu	bu	ca	ca	ca	ca	ce	ch
9	Adtile Technologies Inc.	Adtile Technologies Inc. is a provider of multi-sensor adve	4			1																						
14	Sticar	Sticar is a new digital Out-of-Home platform that help adv	5			1																		1	1			
41	Ebaoyang	Ebaoyang is an online-to-offline car maintenance service c	1																						1			
43	Endado	Buy the auto part your car fits and find a good mechanic.	1																						1			
55	Engie	Engie is a car repair marketplace offering a full solution to	3																						1			
63	My Dealer Service	My Dealer Service empowers automotive service departm	2											1														
72	PARts	PARts is an e-cosystem of tools that helps the auto parts ir	3																									
114	pitstop	Making automotive servicing Safer, quicker, and more affc	1											1														
133	Snapsheet	Snapsheet, an app for smartphones and the web, allows u	3																									
135	TyresOnTheDrive.com	TyresOnTheDrive.com is an online business offering a hug	1																		1							
145	Wheelwell, Inc.	Market network for the automotive parts and services mai	1											1														
155	Who Can Fix My Car	Who Can Fix My Car is an online car repair and servicing ma	1																						1			
161	ABEJA	Super Flexible Artificial Intelligence Platform include Dee	2																									
167	AdasWorks	AdasWorks develops artificial intelligence-based software	3											1											1			
180	Atooma INC	Resonance AI creates connections between apps & device	3											1														
182	BrainChip Inc.	Leading provider of software and hardware accelerated sc	5	1																								
196	DigitalGenius	DigitalGenius brings practical applications of deep learnin	3																									
215	Locus.sh	Locus.sh is an Indian platform that enables enterprises to i	2																									
224	Sherpa	Sherpa is the predictive Personal Assistant with a strong e	2																									
244	Simularity	Artificial Intelligence For The IoT	1																									
245	Cruise Automation	Cruise Automation is a self-driving car company that devel	2											1											1			
262	Nauto, Inc.	AI-Powered Autonomous Vehicle Technology Company	1												1													

Appendix 4 - Determination of (k) in K-means

Appendix 4 shows the number of items per cluster for different cluster sizes. As the K-means algorithm is not able to automatically calculate the number of clusters and because this approach requires a lot of computing power, the number of clusters has been calculated manually. The chosen cluster sizes are 6, 8, 10, 12, 14, 16, 18 ,20.

Cluster	Number of items per cluster
Cluster 0	123
Cluster 1	12
Cluster 2	10
Cluster 3	54
Cluster 4	19
Cluster 5	20

Number of cluster: k = 6

Cluster	Number of items per cluster
Cluster 0	92
Cluster 1	12
Cluster 2	7
Cluster 3	54
Cluster 4	17
Cluster 5	20
Cluster 6	6
Cluster 7	30

Number of cluster: k = 8

Cluster	Number of items per cluster
Cluster 0	78
Cluster 1	15
Cluster 2	17
Cluster 3	13
Cluster 4	23
Cluster 5	26
Cluster 6	7
Cluster 7	10
Cluster 8	28
Cluster 9	21

Number of cluster: k = 10

Cluster	Number of items per cluster
Cluster 0	14
Cluster 1	17
Cluster 2	7
Cluster 3	11
Cluster 4	11
Cluster 5	23
Cluster 6	22
Cluster 7	36
Cluster 8	19
Cluster 9	22
Cluster 10	17
Cluster 11	39

Number of cluster: k = 12

Cluster	Number of items per cluster
Cluster 0	44
Cluster 1	10
Cluster 2	22
Cluster 3	4
Cluster 4	30
Cluster 5	14
Cluster 6	5
Cluster 7	32
Cluster 8	12
Cluster 9	7
Cluster 10	21
Cluster 11	11
Cluster 12	7
Cluster 13	19

Number of cluster: k = 14

Cluster	Number of items per cluster
Cluster 0	22
Cluster 1	8
Cluster 2	15
Cluster 3	11
Cluster 4	10
Cluster 5	11
Cluster 6	7
Cluster 7	4
Cluster 8	20
Cluster 9	9
Cluster 10	24
Cluster 11	5
Cluster 12	39
Cluster 13	22
Cluster 14	11
Cluster 15	20

Number of cluster: k = 16

Cluster	Number of items per cluster
Cluster 0	7
Cluster 1	16
Cluster 2	6
Cluster 3	11
Cluster 4	4
Cluster 5	12
Cluster 6	15
Cluster 7	23
Cluster 8	7
Cluster 9	19
Cluster 10	9
Cluster 11	25
Cluster 12	5
Cluster 13	10
Cluster 14	21
Cluster 15	11
Cluster 16	28
Cluster 17	9

Number of cluster: k = 18

Cluster	Number of items per cluster
Cluster 0	7
Cluster 1	16
Cluster 2	7
Cluster 3	4
Cluster 4	4
Cluster 5	11
Cluster 6	12
Cluster 7	22
Cluster 8	4
Cluster 9	17
Cluster 10	9
Cluster 11	22
Cluster 12	5
Cluster 13	10
Cluster 14	16
Cluster 15	10
Cluster 16	21
Cluster 17	8
Cluster 18	11
Cluster 19	22

Number of cluster: k = 20

Cluster	Number of items per cluster
Cluster 0	22
Cluster 1	10
Cluster 2	9
Cluster 3	4
Cluster 4	18
Cluster 5	12
Cluster 6	1
Cluster 7	9
Cluster 8	4
Cluster 9	6
Cluster 10	10
Cluster 11	4
Cluster 12	3
Cluster 13	6
Cluster 14	4
Cluster 15	3
Cluster 16	4
Cluster 17	2
Cluster 18	4
Cluster 19	6
Cluster 20	5
Cluster 21	8
Cluster 22	4
Cluster 23	16
Cluster 24	2
Cluster 25	14
Cluster 26	28
Cluster 27	20
Cluster 28	5
Cluster 29	5

Number of cluster: k = 30

Appendix 5 - Cluster (C) 20 Keywords (K) 10

Attribute	Cluster 0	Attribute	Cluster 1	Attribute	Cluster 2	Attribute	Cluster 3	Attribute	Cluster 4
context	0.0640	antenna	0.2467	collis	0.1933	connector	0.3844	simul	0.2299
digitalis	0.0554	channel	0.1434	algorithm	0.1284	contact	0.1775	congest	0.1140
testb	0.0548	scatter	0.1072	window	0.1167	wireless	0.1610	intersect	0.1063
learn	0.0512	relay	0.0810	acceler	0.1006	hardwar	0.1255	traffic	0.1006
busi	0.0501	spectrum	0.0641	primit	0.0898	unifi	0.1220	ticket	0.0911
enterpris	0.0457	radio	0.0597	curv	0.0871	linux	0.1053	vissim	0.0825
domain	0.0412	cell	0.0577	obstacl	0.0859	stamp	0.1000	traffic	0.0782
deploy	0.0410	modul	0.0480	slide	0.0771	processor	0.0963	headway	0.0738
organis	0.0406	frequenc	0.0474	planner	0.0718	gnss	0.0962	park	0.0709
constitu	0.0387	multi	0.0470	filter	0.0696	modul	0.0922	request	0.0655

Attribute	Cluster 5	Attribute	Cluster 6	Attribute	Cluster 7	Attribute	Cluster 8	Attribute	Cluster 9
autom	0.3163	ethernet	0.2806	beacon	0.3163	ada	0.0766	gridlock	0.1118
robot	0.1171	intel	0.1442	messag	0.0798	warn	0.0687	compani	0.1010
traffic	0.0735	secur	0.0571	densiti	0.0627	site	0.0643	unman	0.1009
intuit	0.0731	atom	0.0535	peer	0.0580	traffic	0.0609	transport	0.0966
session	0.0692	flexray	0.0496	authent	0.0574	lane	0.0499	polici	0.0850
fhwa	0.0685	sensor	0.0472	park	0.0558	autonom	0.0487	stake	0.0799
freeway	0.0563	bandwidth	0.0471	local	0.0535	simul	0.0484	railway	0.0749
valid	0.0542	video	0.0456	departur	0.0513	collis	0.0454	emiss	0.0737
slip	0.0511	audio	0.0444	wireless	0.0503	car	0.0412	privat	0.0719
reilhac	0.0504	chip	0.0444	estim	0.0486	situat	0.0405	undergo	0.0701

Attribute	Cluster 10	Attribute	Cluster 11	Attribute	Cluster 12	Attribute	Cluster 13	Attribute	Cluster 14
carshar	0.4431	secur	0.2177	cloud	0.1818	leon	0.1575	packet	0.3162
revenu	0.1065	attack	0.1681	hadoop	0.1738	sync	0.1524	node	0.2137
share	0.1021	cyber	0.0752	song	0.1170	buyer	0.1423	rout	0.1684
profit	0.0907	threat	0.0629	cluster	0.0969	compass	0.1323	protocol	0.1496
taxi	0.0882	authent	0.0623	safeguard	0.0814	onstar	0.1254	junction	0.0867
sale	0.0860	protect	0.0494	diagnost	0.0760	standstil	0.0940	deliveri	0.0758
citi	0.0805	privaci	0.0483	data	0.0733	autonom	0.0889	relay	0.0663
retriev	0.0778	polici	0.0426	analyt	0.0733	campaign	0.0888	predict	0.0661
premium	0.0747	firmwar	0.0414	lake	0.0568	advertis	0.0799	metric	0.0626
hail	0.0697	certif	0.0411	databas	0.0565	ecosystem	0.0781	overhead	0.0574

Attribute	Cluster 15	Attribute	Cluster 16	Attribute	Cluster 17	Attribute	Cluster 18	Attribute	Cluster 19
wheeler	0.2673	charg	0.1090	interior	0.2229	display	0.0954	weather	0.0744
semiconductor	0.1635	energi	0.1061	concept	0.0980	motorist	0.0724	patent	0.0655
powertrain	0.1157	batteri	0.0872	seat	0.0960	sign	0.0699	traffic	0.0570
telemat	0.0818	electr	0.0843	tour	0.0731	touch	0.0662	traffic	0.0547
govern	0.0753	bev	0.0776	citi	0.0729	cloud	0.0580	ecosystem	0.0527
electron	0.0677	grid	0.0558	disrupt	0.0655	traffic	0.0578	simul	0.0525
buse	0.0671	voltag	0.0503	peopl	0.0548	chapter	0.0527	digit	0.0510
session	0.0654	har	0.0501	autonom	0.0533	navig	0.0511	hazard	0.0504
compani	0.0601	servic	0.0458	space	0.0515	infotain	0.0495	pattern	0.0473
adopt	0.0563	stop	0.0450	social	0.0456	screen	0.0478	smart	0.0467