Master of Science
in
Business Administration

Entrepreneurship, Innovation & Strategy

Discovering technologies in the automotive industry

Aaron Sandberg

24 February 2024

1st Examiner: Dr. Igors Skute

2nd Examiner: Dr. Matthias de Visser

UNIVERSITY OF TWENTE.

Abstract

Companies increasingly face volatile, uncertain, complex and ambiguous environments. Firms need to adapt to the environment to stay competitive. To achieve this, companies invest in Research and Development. But firms have limited resources to spend, so the main problem is selecting the best innovations to pursue. To gain insights into future requirements, companies engage in Technological Foresight. This research proposes using text mining on patents to extract information for Technological Foresight. More specifically, the study intends to discover current technologies in the automotive industry, classify them based on uncertainty characteristics, and create an innovation portfolio uncertainty matrix. Topic Modeling was used on patent titles to uncover topics. These topics were then assigned uncertainty scores and plotted in the uncertainty matrix. The technologies discovered are diverse and vary in uncertainty, with advanced driver assistance systems and alternative power sources being the most uncertain ones. The uncertainty matrix shows the technologies' distribution to be skewed towards low uncertainty. According to theory, the distribution should be optimized to include more uncertain technologies. It is discussed that this stipulation might not apply to the automotive industry to the same extent, due to industry-specific characteristics. This research demonstrated how text mining can be used in practice to discover insights.

Table of Contents

1. Introduction

Companies are increasingly facing volatile, uncertain, complex and ambiguous (VUCA) environments (Mack & Khare, 2015). These environments create a need for companies to become more flexible and dynamic in their activities, to deal with future challenges better (Bennett & Lemoine, 2014). Industries have resultingly become more fast-paced, but environmental pressures are still increasing. This is partly due to policymakers around the world tightening regulations such as to fight climate change, which means that companies must lower emissions and become more sustainable (European Environment Agency, 2016; Lang & Murphy, 2014). Additionally, digitalization and the continuing evolution of technology make companies change the ways they operate. Because of facing these environmental pressures, firms need to adapt to the pressures faced, to remain competitive (Allen et al., 2017).

Research and Development (R&D) enables companies to adapt to these pressures, which makes it an important function, especially in High-Tech industries (Mallick & Schroeder, 2005). By investing in R&D, companies try to adapt to the pressures faced. The problem with this is that R&D Management is a complex, resource-consuming and uncertain discipline (Eckert & Hüsig, 2021). The influence of R&D on the company can be existential, with successful innovations benefitting the company, and disruptions by competitors threatening the companies' survival (Mitchell & Hamilton, 1988). Due to limited resources, companies cannot invest into every opportunity, which means the existing resources need to be allocated optimally. Following this, the need for optimal portfolio management arises.

R&D is essential in the automotive industry, as it is characterized by complex technologies in a highly competitive environment (Mallick & Schroeder, 2005). A slogan of the automotive company Audi AG reflects this, namely 'Vorsprung durch Technik', which can be translated to 'Leading edge through technology' (Audi AG, 2021). Annually, $116B are spent on R&D in this industry (R&D World, 2022). This is due to the industry's size, with $2.86T annual revenue (IBIS World, 2022). Automotive companies need to adapt to rising external pressures to be successful globally. To achieve this, R&D management is used to adapt and improve their product offering (Iamratanakul et al., 2009). Due to the size of the companies and their resulting investments in R&D, there are many innovation projects pursued all at once, requiring many of their limited resources. Shortly put, the main problem automotive companies face is selecting the best innovations to pursue.

The automotive industry faces some external challenges, with one of the main challenges being sustainability (Wellbrock et al., 2020). Climate change is driven by carbon emissions, partly by cars, which additionally worsen air quality (Qian et al., 2000). Additionally, demands for components, such as secondary batteries to become more ecologically friendly are raised by policymakers (Vandepaer et al., 2017). On a different note, pressure and uncertainty are created through automotive companies competing against each other to develop new successful technologies.

Cars have become extremely advanced technologically over the last 50 years (Taub et al., 2007). Exemplary of this, electronics have become more complex, as digital systems control cars and their many subsystems (Llopis-Albert et al., 2021). This complexity requires much knowledge in different domains, making the coordination and selection of the innovation project portfolio more complicated. Adding to this, the previously mentioned limited resources add to the importance of selecting the best innovation projects. To gain more insights into what the future might require, companies engage in activities like discovering trends and new or leading technologies. This is activity is often called Strategic Foresight (Pietrobelli & Puppato, 2016).

There are large ever-growing amounts of textual data available, which can generate insights (Pope et al., 2000). This data is often relatively easy to access as well, when contrasted with other methods. This makes text mining a useful tool for Strategic Foresight (Kayser & Blind, 2017a). Over the last two decades, more and more algorithms have been used to find patterns in large datasets. Text mining can analyze large amounts of data in short timespans. It is an extremely cost-efficient technique compared to labor and may generate new insights trough identifying previously unknown patterns and developments (Antons et al., 2020). In this research, text mining is used on patents with the intention to extract information aiding to foresight technologies.

Through text mining, the data is prepared as an input into a theoretical framework, which classifies technologies into different categories according to their respective uncertainty characteristics. This results in a visualization of the technology portfolio in the automotive industry measured on two types of uncertainties.

This research closes the research gap of how using text mining on patent data can be used to construct the uncertainty matrix.

This research explores current technology themes in the automotive industry, by applying text mining to patents. Additionally, this research uses the discovered themes to assign them to companies, after which an innovation portfolio is created. The objective is to advance our understanding if and how text mining in combination with innovation portfolio techniques can lead to a useful clustering of automotive companies' patents in the portfolio matrix, which reflects the automotive industry. This approach can further provide indications on the feasibility of using patents to gauge innovation portfolios, as well as the degree to which different companies in the same industry can be distinguished through this approach. Thus follow the research questions:

How can unsupervised text mining uncover current technology themes in the automotive industry?

What are the current technologies in the automotive industry?

How do the individual portfolio compositions of the leading companies vary?

This research contributes academically to the domain of R&D management and provides insights by applying the theoretical framework derived from MacMillan & McGrath (2002) to patents. In scientific literature, there is just one other paper using the same framework. Luo (2011) stochastically describes market and technological uncertainty. Thus, there is a lack of research applying the framework, which this paper aims to improve. Additionally, it builds upon the evidence that text mining can be a valuable tool in the domains of R&D management and technology foresight.

Regarding practical implications and relevance for practitioners, this paper demonstrates how text mining can be used to identify themes from patent data in the automotive industry, which may be a viable method in practice to use for detection of technology themes and comparison of companies innovation portfolios. Through clustering into a technological/market uncertainty matrix framework, early judgements regarding the diversification of innovation portfolios in the automotive industry are made possible. On another note, this research demonstrates how the R&D project portfolio matrix can be applied to patent data. Additionally, it discusses how well the theoretical stipulations regarding portfolio distribution match the ability of the automotive industry to fulfill them.

Policymakers may derive value from this research through using the methods themselves to track technologies, which with regards to climate change can serve as an indicator which technologies the innovation focus seems to be on. Armed with this knowledge, policymakers can roadmap the automotive industry's progress and adjust legislation.

2. Literature Review

2.1 R&D Management

R&D Management can be defined in many ways, but one commonly used definition is the OECD's definition: "Research and development (R&D) comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge (including knowledge of man, culture and society) and the use of this knowledge to devise new applications. R&D covers three activities: basic research, applied research, and experimental development." (OECD Publishing, 2012). In high-tech industries, where technological change is constant, R&D is a very important department of the company. Especially in the automotive industry, there is a high rate of technological innovation due to large investments in R&D, especially since the 21st century (Gerhard et al., 2008). As R&D develops the products and services of the company, it is a vital function to ensure future success by introducing superior innovations and thus at least keeping up with competing firms' value propositions (Yalcinkaya et al., 2007). Especially in fast-paced industries, such as the automotive industry, companies need to invest heavily in R&D to not be left behind (Madhok & Osegowitsch, 2000). When the rate of change is high, there is high uncertainty regarding the future. Jalonen (2011) identifies the main sources of environmental uncertainty in the R&D process: Technological uncertainty relates to which technologies will displace the current technologies. Market uncertainty describes the uncertainty whether the product will be adopted in the market. Regulatory uncertainty describes the possibility of new laws which may restrict or subsidize an industry or technologies in some way. In the automotive industry, development cycles are long and require large investments. This makes companies less agile and able to react quickly to changes. Thus, the need for optimized innovation portfolios arises.

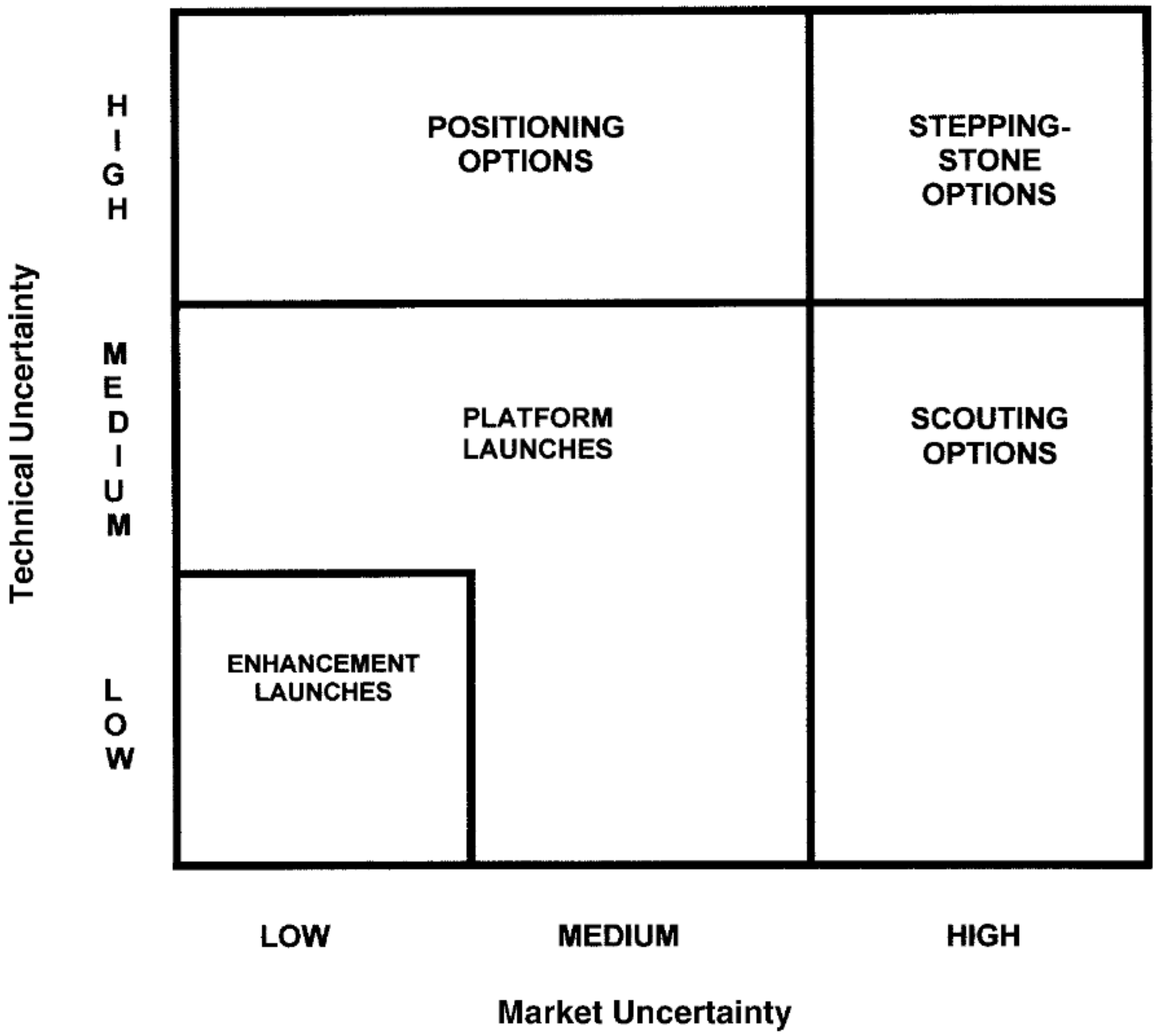
To successfully innovate long-term, portfolio management techniques are used in the R&D domain. Portfolio management is the strategic process of managing the innovations a company is researching and developing. It tries to select innovations congruent with the company's strategy to provide fitting future revenue streams (Jugend & Da Silva, 2014). Further, and most importantly, it allocates the limited available resources of the company to the projects that seem most attractive. In line with this, it balances the number of projects, and aims to strike a balance between more and less radical innovations (Cooper et al., 1999). An innovation portfolio can result in path dependency, which is one factor for sustaining competitive advantage long term (Reuter et al., 2010). This portfolio needs to balance between incremental innovations that enhance the product, and innovations which are more uncertain, but may shape the future of the companies.

The technological/market uncertainty matrix framework by MacMillan & McGrath (2002) is the main framework used in this paper. It looks at innovative ideas with a real options mindset, similar to the stock market. An option grants the right to exercise the option (project) in the future given favorable developments and does not have to be used. Applied to R&D project portfolios, using this method makes most sense given highly uncertain but promising projects, to probe the projects with minimal investment and gain knowledge. The framework categorizes projects into different categories based on the degree of the respective uncertainty, which can help a company assess how their project portfolio's distribution looks put onto this matrix. In this case, technological and market uncertainty are used as indicators of innovation uncertainty. The technological and market-based views two provide two different perspectives of an innovation project. The technological view focuses on the innovation itself and how the innovation works. The market-based view is focused on the customer, not the product. Furthermore, it is about what markets the innovation could be successful in, and thus what customer segments to target. The framework used adds the degree of uncertainty to these two views, which results in two distinct variables which measure the innovation's uncertainty.

The distribution of the portfolio gives an indication how well balanced the portfolio is, as it helps companies optimally innovate by neglecting neither types of projects. Additionally, the portfolio needs to fit the company's technology strategy as well as general strategy. Different industries require different foci: a high-tech company with high technological uncertainty will focus more on not missing emerging technologies by using many real options. A cement manufacturing company may focus more on new market explorations and improving the current product.

Figure 1

The Uncertainty Matrix by MacMillan & McGrath (2002)



The different types of options in this framework are determined based on the position in the matrix. This matrix consists of technological uncertainty on the y-axis and market uncertainty on the other, as seen in *Figure 1*. Positioning options are made up from high technological uncertainty and at most medium market uncertainty. In this case, the technology trajectory is unknown, there may be many technologies but no dominant one yet the market is more or less known. The most important activities for an innovation in this category are knowledge acquisition through experiments and making many small bets which technology will become dominant. Scouting options exhibit high market uncertainty, whereas technological uncertainty is not larger than medium. As the name suggests, scouting options are used to enter new markets or segments with a clear vision of the technology used. The main concern is finding target markets and segments, as well as understanding customer needs. Resultingly the main activities are knowledge acquisition and experimenting through user feedback. Stepping-stone options are a combination of high uncertainty in both technology and markets. These are key for a company's long-term strategy to evolve. These options are highly attractive for the company, and have somehow been noticed, but since neither the technology nor market are clear, a lot of exploration is necessary. To start, a company can start with related but less demanding markets, to improve their technological expertise for the technologies to be developed and after reducing the technological uncertainty progress to the target market. Platform launches and enhancement launches are not suited for the options approach, as they are technologically and market-wise certain enough that other techniques or going straight to development are more effective. Yet for an assessment of the R&D portfolio distribution, these categories also need to exist. Platform launches are at most of medium uncertainty and are often the next generation of a product such as a new drivetrain platform for a car. They are more complex than enhancement launches and form the basis for medium-term competitive advantage. Enhancement launches are small variations or improvements to existing products, with a clear understanding of technology and market from the outset.

2.2 Technology Foresight

Technology foresight is a term derived from foresight, which Tegart (2014) defined simply as "Foresight is a systematic approach to understanding and engaging the future". Technology foresight thus concerns understanding the technologies of the future. Especially since the advent of computers, industries have become more difficult to survive in for companies. Reasons are the decreasing lengths of product life cycles, and quicker time-to-market of innovations, the increased crossovers of technologies into other industries, globalization of markets and higher risks of disruption by missing the chance to jump on new technologies (Reger, 2001). These characteristics apply especially to high-tech industries, which includes the automotive industry studied in this paper. The concept of technology foresight pairs well with the uncertainty matrix by McMillan & McGrath (2002) used in this paper, as the frameworks intention to build a diversified portfolio in terms of uncertainty of innovations is exactly what technology foresight prescribes: An approach that explores many types of options to be prepared for future developments with probing for information by treating projects as real options. This paper relates to technological foresight as well, as mining patents is a way to detect current but rising technologies which may shape the future.

The developments of global markets becoming more uncertain (VUCA) is both a constant threat but also opportunities for companies. It could be a threat because the company has to take steps to adapt to steer through the uncertainty, which calls for useful approaches to prepare for the future. Through preparing thoroughly for the future, these developments can transform from being a threat to an opportunity, as good technology foresight can give a company a leading edge over competitors. Technology foresight has multiple notable characteristics (Pietrobelli & Puppato, 2016): Firstly, foresight means trying to guess what the future holds. It can steer technologies into the envisioned direction and make itself self-fulfilling (Miles, 2010). Secondly, it often is participatory and a broader range of stakeholders can input their views, leading to more sophisticated strategies and network creation and collaboration. Technology foresight is not necessarily focused on companies alone, but often works together with universities, governments and society. Lastly, technology foresight exists at different levels, from organizational and regional, to national and supranational. Especially in developing countries, governments should collaborate with industries on technological foresight to accelerate the development of the country, as exemplified by South Korea, which experienced strong growth and is now one of the most technologically knowledgeable countries (Pietrobelli & Puppato, 2016). This makes clear that technological foresight is not just an activity for companies alone.

The threat of becoming uncompetitive through missing the signals in the environment about disruptions is mitigated by technology foresight (Pietrobelli & Puppato, 2016). The case of Nokia exemplifies what can happen in the worst case when missing the emergence of a market change. From dominating the phone market as largest producer by far, they were disrupted by Apple and Samsung through superior technology and understanding of customer needs. This led to their downfall in just a few years. The reason was mainly a lack of understanding their customers. Nokia did not understand their customers well enough, which made them incapable of foreseeing new technologies, as technologies become successful through serving the customer better. It becomes clear how technology foresight is not a standalone activity but embedded in the organization and dependent upon the PESTEL environment, not just the technological landscape (Bouwman et al., 2014; Junquera et al., 2016; Tegart, 2014).

The field of technology foresight can be divided into two general types of methods: qualitative and quantitative approaches (Magruk, 2011). Qualitative methods are mainly based on expert knowledge, whereas quantitative methods are based on data. A more advanced distinction is provided by Porter (2010), who reviews the eleven major approach families used in the domain of technology foresight. Magruk (2011) provides an even more comprehensive overview with more than 120 specific methods mentioned. The large number of methods is due to the breadth of the domain of technology foresight, according to the author.

Firstly, there are creativity approaches, which try to employ creativity to arrive at new insights. A common method in this methodological family is TRIZ, which is the theory of inventive problem solving (Ilevbare et al., 2013). Next are monitoring and intelligence approaches, which are often based on data mining (de Miranda Santo et al., 2006). Descriptive approaches follow, which include popular methods such as bibliometric analysis and impact analysis (Minghui et al., 2022). Another method family are matrices, such as the uncertainty matrix used in this paper (MacMillan & McGrath, 2002). Statistical analysis is another method family. Trend analysis is used to visualize the changes in popularity of the subject of analysis, and one of the most common methods used in this domain (Magruk, 2011). Expert opinion approaches such as the Delphi method, surveys, focus groups and other participatory approaches are another common qualitative method family (Aichholzer, 2009; Kanama et al., 2008). Next come modelling and simulation approaches, which use techniques such as hybrid simulation and exploratory modeling (Kolominsky-Rabas et al., 2015; Kwakkel & Pruyt, 2013). Then comes logical or causal analysis, which comprises requirement analysis, stakeholder analysis and policy assessment among others (Apreda et al., 2016). Then follows one of the most popular approaches, namely roadmaps, which is used to roadmap technologies and products and science (Hussain et al., 2017). Then follow scenario approaches, which are also often combined with roadmaps and one of the most common as well (Drew, 2006; Hussain et al., 2017).

Having discussed the major families of approaches, the state-of-the art is dominated by three major methods: the Delphi method, scenario analysis and roadmaps according to Minghui et al. (2022). These authors conducted an extensive literature review which is further summarized in this paragraph. Generally, more qualitative methods than quantitative methods are used in the domain of technology foresight, but there seems to be a trend towards combining the methods. More technologically advanced methods such as text mining and data analysis are increasingly gaining in popularity in this field. It is stipulated that qualitative studies based on experts are difficult to reproduce as they are mainly based on implicit knowledge, making them less valuable to others seeking to use technology foresight in practice. Furthermore, there is a development towards using multiple methods at once, to gain a more comprehensive and reliable understanding compared to using a single method only.

Further in their literature review, the authors found that technology foresight is most often researched on the enterprise level, with governmental and organizational level following in that order. Strategic Foresight is the most popular research topic at the moment, with the focus of it being mainly on the enterprise level, whereas the higher levels such as national levels seem to be out of focus. While there is much research on theory and practical application, there is a lack of literature review papers. This lack of literature reviews makes it difficult to capture the state of the art, as identifying articles that advance the field requires extensive knowledge on what has already been studied in the field (Kraus et al., 2022). This means that it is challenging for researchers who are not experts in the field to discern these articles from others (Palmatier et al., 2018). Another lack is apparent in the field of detecting and evaluating technological foresight, which this research aims to bridge (Minghui et al., 2022). In addition, patent analysis is one of the less common methods used for technology foresight, which this research also aims to improve.

Cifci & Yuksel (2018) forecast the sixth generation of technology foresight, which will be concerned with the 'netocracy', a word combining internet and aristocracy. This netocracy will use big data, artificial intelligence, machine learning and cyberware, which are forecasted to become dominant technologies and resultingly be widely adopted in management and technology. Thus, there will be a change in methods in the domain of technology foresight, with more methods harnessing big data and digital data in general.

Murry & Hammons (1995) stipulate that traditional technology foresight approaches are resource-consuming, such as the Delphi method taking a small number of years at times and requiring many experts to function. Additionally, experts introduce a relatively large subjective bias, which research should seek to mitigate (Takahashi et al., 2014). This research fits into the domain of technology foresight by employing a less-utilized approach to demonstrate its viability, while mitigating the problems associated with traditional and expert-based qualitative methods.

2.3 Text Mining in R&D Management

Text mining is a technique applied on textual data to extract information. It can process large amounts of data in a short time (Wang et al., 2010). Further, parts of the process can be automated and require less human input. For application of text mining in R&D management, relevant data needs to be available and fitting methods used to obtain meaningful results. There are many relevant types of data sources available. Research papers and patents are some of the most frequently used data sources (Porter, 2015). Others are grant applications, new product announcements, newspaper articles, customer reviews and social media (Antons et al., 2020).

Text mining encompasses a few kinds of commonly used techniques. There are two main technique types used: The first are dictionary-based techniques, which use a predefined list of words to scan the textual data. The second are algorithms, which can be divided into supervised and unsupervised execution. Supervised algorithms require human input, such as discovering and creating categories to sort into by manually classifying a training dataset and applying it to a testing dataset afterwards. Classifying techniques range from binary classification to multi-classification. Unsupervised algorithms do not need human input and cluster instances based on similarity or distance indicators. Labeling of clusters needs to be done afterwards, as the algorithm does not 'know' the appropriate label for a discovered cluster. Clustering has become the most used method in R&D text mining and disproportionately gained share in recent years compared to the other mentioned methods (Antons et al., 2020). In the domain of R&D, text mining was found to be frequently used for technological forecasting and technology management. Creating roadmaps of technologies was another frequent intention of use (Antons et al., 2020). Text mining has been frequently applied to the automotive industry as well, such as by Stoehr et al. (2020), who performed a network analysis to analyze corporate positioning and trending technologies.

Table 1

Literature Review of Text mining in the R&D Management Domain

Author	Data Type	Algorithm Type	Result
de Miranda Santo et al. (2006)	Scientific Articles	Keyword-Filtering	Identified keywords in nanotechnology
Lee et al. (2020)	R&D Project data	Frequency analysis, association rule mining, topic modeling	Discovered keywords, associations and topics in sustainable graphene R&D projects
Wang et al. (2010)	Patents	KeyGraph	Discovered technology scenarios and used TRIZ to discover evolutionary patterns
Antons et al. (2020)	Scientific Articles	Classification, Clustering, Dictionary-based	Reviews the use of text mining in scientific R&D literature
Hu et al. (2022)	Patents	Topic Modeling	Identified topics and predicted trends using a model
Lee & Kang (2018)	Scientific Articles	Topic Modeling	Identified topics of a literature field
Park et al. (2018)	Patents	Topic Modeling	Performed trend analysis
Tian et al. (2022)	Patents	Topic Modeling	Identified topics
Riesener et al. (2022)	Scientific Articles, Social Media Content	Topic Modeling	Identifies product innovation potentials
Shih et al. (2010)	Patents	Patent Trend Change Mining	Monitors trend changes

In Table 1, a literature review of text mining in the domain of R&D management is found. One simple approach to gain knowledge from large datasets is to use keyword filtering. De Miranda Santo et al. (2006) identified keywords in the nanotechnology field using this approach. Lee et al. (2020) used a simple frequency analysis to discover keywords and other methods to demonstrate how knowledge can be uncovered about an industry.

Moving to more complex methods, Wang et al. (2010) used an algorithm to discover technology scenarios and adopted a theoretical framework for innovation (TRIZ) to find evolutionary patterns based on this. This shows how theoretical frameworks can be operationalized through text mining. Text mining can convert data into a fitting input for theoretical frameworks, as Hu et al. (2022) show by using topic modeling to construct a ARIMA model.

The topic modeling approach used in this research is similar to other research papers such as Lee & Kang (2018), who used the same algorithm to discover core topics in the domain of technology and innovation management. Topic Modeling is a commonly used method to analyze patents and scientific papers (Antons et al., 2020) and is used for various objectives in the R&D domain. Tian et al. (2022) identified industry-specific technological topics from patents using this approach. Another study identified product innovation potentials from scientific articles and social media content (Riesener et al., 2022). Similarly related to foresighting, Park et al. (2018) employed topic modeling to create a trend analysis for Artificial Intelligence patents. Other text mining methods may also enable trend change monitoring (Shih et al., 2010).

Generally, text mining in the R&D and Technology foresight domain is often used to identify relevant topics, trends and their change, and as input for theoretical frameworks or models. It is also used to identify incumbent and emerging rivals and monitor rival's activities and progress (Kehoe & Yu, 2001).

More specifically, the methods used in this paper belong to the domain of patent mining. The first papers applying such techniques emerged around the 2000's (Kehoe & Yu, 2001; Larreina & Hernando, 2005). Before that, patent analysis without text mining, was already an established topic and had been applied for decades (Neudorfer & Voos, 1977). Patent mining has evolved further form its inception, with more refined methods and scientific papers published (Madani & Weber, 2016).

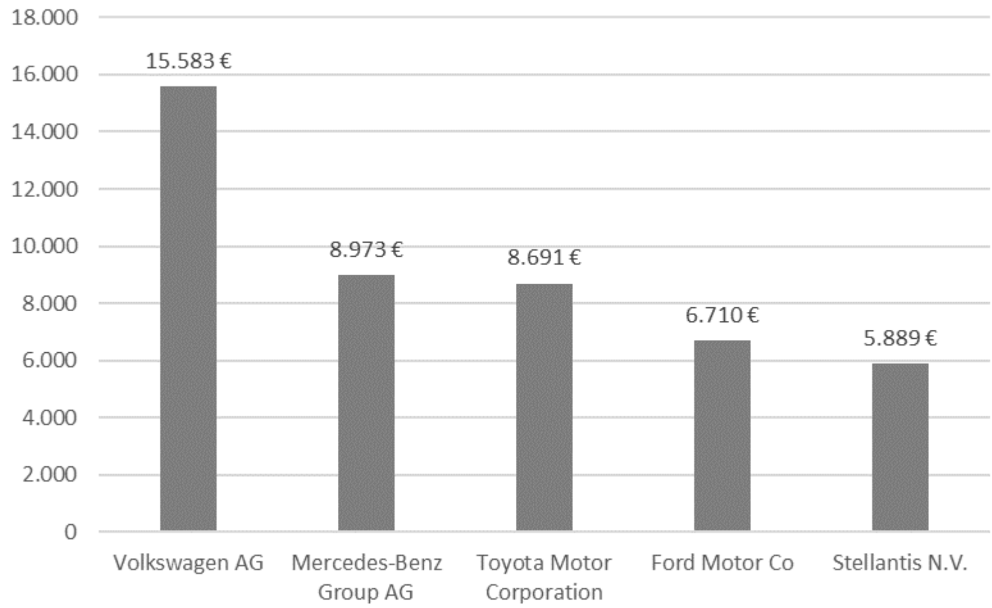
3. Methodology

3.1 Research Context

This section explains how automotive companies were selected and why they are relevant in this research context. Automotive companies were selected because of large investments in R&D, external pressures to innovate. The automotive industry is one of the largest industries in the world, with the global automotive industry valued at 2.52 trillion U.S dollars in 2022 (IbisWorld, 2023). The five companies with the most revenue were selected. The following figure displays their research & development expenses in million U.S dollars from 2021 (European Commission, 2022).

Figure 2

Spending in Million U.S. Dollars of the Selected Companies in 2021



Note. Source: European Commission (2022)

Together, the firms accumulated 45.8 billion U.S dollars in R&D expenses in 2021, which makes clear that innovation is important for automotive companies. With legal and societal pressures mounting, automotive companies are focusing intensely on R&D to address especially sustainability problems among others. The 2022 annual report of Volkswagen AG highlights this, with Volkswagen AG having increased their R&D spending by 21.3%, focusing mainly on electric vehicles, digitalization and high-tech systems (Volkswagen AG, 2023). Not all automotive companies were selected, since part of the discussion looks at the individual company level, which is unviable to perform for all companies in the industry.

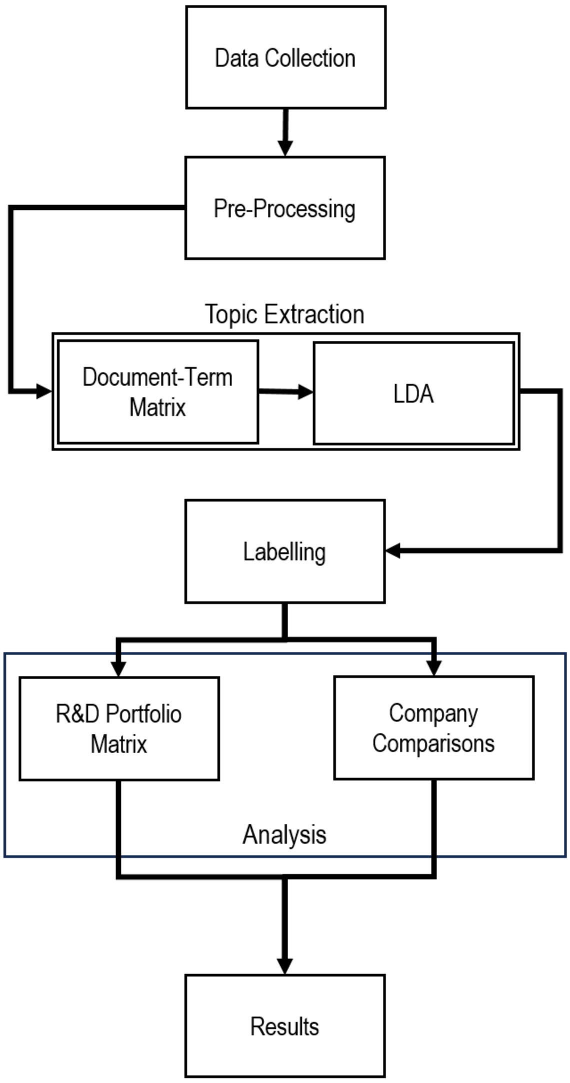
Patents are the units of observation in this research. Patents were chosen, as they are written in a standardized format, are readily available & veracious. Patents also present shortcomings, such as data noise which needs to be removed (Kim et al., 2019). Additionally, they only contain patented technologies, which not every innovation is. Nevertheless, patents were chosen, due to being well suited for this research.

3.2 Methodological Process

The objective of this research is to utilize advanced text mining techniques on patent titles to capture and analyze emerging technologies within the automotive industry. It is additionally outlined in Figure 3. For transparency and replication purposes, the code used is provided in Appendix 6.3 (Woodside, 2016). The dataset cannot be provided due to license limitations but may be found elsewhere.

Figure 3

Methodological Process



3.3 Data Collection

The dataset used was retrieved from Orbis, which is a database with detailed information on companies. Orbis was created by Bureau van Dijk, a Moody's company (Bureau van Dijk, 2023). The industry classification standard NACE Rev2 of the European Union(Office for Official Publications of the European Communities, 2008) was used to identify & download the patent tiles of the five companies with the most annual revenue in 2022. The companies chosen were Volkswagen AG with $311B, Toyota Motor Corporation with $278B, Stellantis N.V. with $192B, Mercedes-Benz Group AG with $164B and lastly Ford Motor Company with $158B. Drilling down on the exact selection of patents, only patent titles were chosen, not full descriptions. This is due to full descriptions not being accessible from the available sources. Nevertheless, patent titles reflect the patent, and can be used to generate relevant topics. Only active granted patents were chosen. Expired patents were excluded due to being too old to be relevant, or not deemed relevant anymore by the company itself. Pending patents were not included, as they have not been granted which would confirm that they are unique.

3.4 Pre-Processing

Due to its functionality being optimally suited to data mining, R was used to import, pre-process, clean and analyze the dataset. R is a programming language for statistical computing and graphics. It allows users to create extensions, also called packages, which add functionality, such as the package used to run the Latent Dirichlet Allocation. The list of packages used can be found in Appendix 6.4

The dataset needs to be pre-processed for the LDA to produce the best results (Haddi et al., 2013; Kim et al., 2019). This involves tokenization, which splits sentences into words. Afterwards, data noise was removed, such as punctuation, numbers, symbols, separators, split hyphens, and non-Latin characters (Tong & Zhang, 2016). Additionally, all letters were transformed to lowercase, and stemmed, which means transforming each word to its wordstem. Stopwords were also removed, such as prepositions, which detract value from the results. Irrelevant words specific to the data, such as 'method' were also removed, for the same reason. Bigrams were created and the corpus was transformed into a document-feature matrix. The matrix was trimmed, to focus on common but distinct terms. Only the top 20% most frequent terms which are present in less than 10% of all documents were kept. The parameters were tested against other parameter configurations, but the aforementioned ones yielded the best results.

3.5 Latent Dirichlet Allocation (LDA)

In this research, Latent Dirichlet Allocation (LDA) with Gibbs sampling was used to uncover latent thematic structures within the patent titles. The maximum number of iterations was 2000, alpha was set at 0.5 and beta at 0.1. The algorithm was developed by Blei, Ng, and Jordan (2003) and has become a popular method due to its ability to analyze large volumes of textual data and reveal latent structures without supervision. It has been applied to various business contexts, such as market research and competitive intelligence, but also academia (Moro et al., 2015; Vílchez-Román et al., 2019; W. Wang et al., 2018). Shortly put, LDA is an algorithm well suited to the task of discovering emerging technology themes form the patent titles dataset.

The following equation and individual explanations of its components are directly cited from Jelodar et al., 2019, pp. 5–6.

LDA assumes that each document can be represented as a probabilistic distribution over latent topics, and that topic distribution in all documents share a common Dirichlet prior. Each latent topic in the LDA model is also represented as a probabilistic distribution over words and the word distributions of topics share a common Dirichlet prior as well. Given a corpus D consisting of M documents, with document d having N_d words ($d \in 1,..., M$), LDA models D according to the following generative process [4]:

Choose a multinomial distribution ϕ_t for topic t ($t \in \{1,..., T\}$) from a Dirichlet distribution with parameter β

Choose a multinomial distribution θ_d for document d ($d \in \{1,..., M\}$) from a Dirichlet distribution with parameter α. (c) For a word w_n ($n \in \{1,..., N_d\}$) in document d,

i. Select a topic z_n from θ_d.

ii. Select a word w_n from ϕ_{zn}.

In above generative process, words in documents are only observed variables while others are latent variables (ϕ and θ) and hyper parameters (α and β). In order to infer the latent variables and hyper parameters, the probability of observed data D is computed and maximized as follows:

$$p(D|\alpha,\beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha)\left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(Z_{dn}|\theta_d)p(w_{dn}|z_{dn},\beta)\right)d\theta_d \qquad (1)$$

Defined α parameters of topic Dirichlet prior and the distribution of words over topics, which, drawn from the Dirichlet distribution, given β. Defined, T is the number of topics, M as number of documents; N is the size of the vocabulary. The Dirichlet multinomial pair for the corpus-level topic distributions, considered as (α, θ). The Dirichlet-multinomial pair for topic-word distributions, given (β, ϕ). The variables θ_d are document-level variables, sampled when per document. z_{dn} , w_{dn} variables are word-level variables and are sampled when for each word in each text-document.

3.6 R&D Innovation Portfolio Matrix

This section discusses the matrix for evaluating R&D portfolios by MacMillan & McGrath (2002). The R&D innovation portfolio matrix is used as it provides a theoretical framework for plotting and evaluating the technological and market uncertainty of technologies. This serves as an indicator for the radicality of technologies. This indication is used to draw conclusions on the automotive industry and to compare it to the theory's prescribed portfolio distribution.

The framework is operationalized through the items in scoring the technologies. This means that the questions originally proposed within the framework are used. The scoring is done as objectively as possible. To make the scoring less subjectively biased, information about the automotive industry and its technologies is gathered, which gives the researcher an improved understanding backed by official information. The information sources used are annual reports of the analyzed companies, industry reports and news articles.

Moving to the matrix itself, the x-axis plots the degree of market uncertainty, which is computed based on the items of the sample questions put forward by MacMillan & McGrath (2002). The degree of uncertainty is operationalized through a scale from 1 to 7, with 1 signifying certainty about the question's answer and 7 meaning high uncertainty about the answer. An example follows: "How certain is your team of the following: The market demand for future products using the fruits of the project." The items only assess how certain the respondent is of the market demand, however high it may be. Thus, it does only measure uncertainty, not potential. The same applies to technical uncertainty, which is plotted on the y-axis. Low uncertainty is defined as below three, medium uncertainty as between three and five, and high uncertainty as above 5. Lastly, the matrix contains the types of innovation projects with prescribed percentages for a strategic, balanced innovation portfolio. The questionnaires may be found in Appendix 6.2 as well as the uncertainty scores.

4. Results

4.1 Descriptive Statistics

All patents of the five automotive companies were downloaded from the ORBIS database. In total, 1.068.719 patent titles were retrieved. After filtering out inactive and not granted patents, duplicates, as well as patents older than 2020, 11.334 patents remain. A sample of patent titles before pre-processing is found in Table 2.

Table 2

Patent Title Samples before Processing

Patent title	Status	Publication	IPC Main	Current owner
Bifurcated air induction system for turbocharged engines	Granted	US11261767B2 On 10/11/2020	F01M13/02	FCA US LLC, MICHIGAN
Car rear hatchback door panel	Granted	CN306963730S On 13/08/2021		Peugeot Citroen Automobiles SA
Method for operating driver assistance system and driver assistance system	Granted	CN109311480B On 27/02/2017	B60W30/18	Volkswagen Aktiengesellschaft

In Figure 4 below, the distribution of patents granted over the years 2020 to 2023 is displayed. The trend of a reduced number of patents granted per year is visible. Though this development may possibly be caused by the time lag between the date of the patent filing, and the date of the patent grant (World Intellectual Property Organization, n.d.). Additionally, Covid-19 affected the automotive industry in several ways and could have resulted in less patent filings (Klein et al., 2021; Nayak et al., 2021).

Figure 3

Distribution of Patents by Year from 2020 to 2023

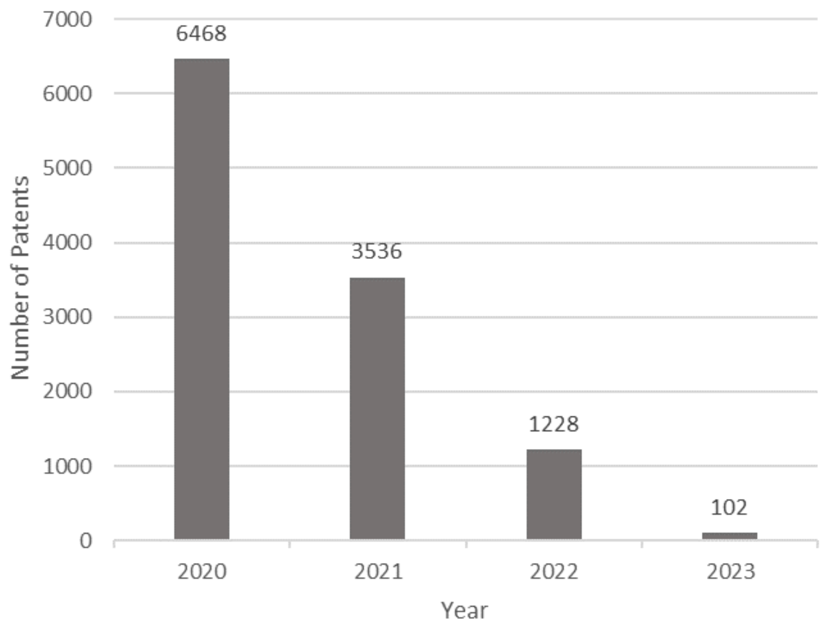


Table 3 below shows the distribution of patents by company. It shows that Toyota Motor Corporation is responsible for approximately 46% of all patents in the dataset. Opposingly, Mercedes-Benz Group AG has the lowest number of patents in a stark contrast to the other companies with 2,49% of the total number of patents. Volkswagen AG, Ford Motor Company and Stellantis N.V seem to be somewhat comparable in patent numbers despite contrasting patent numbers, considering the outliers of Toyota Motor Corporation dominating the dataset and Merces-Benz Group AG being a marginal contender.

Table 3

Distribution of Patents by Company

	Number of Patents	**Share of Patents**
Toyota Motor Corporation	5.265	46,45%
Ford Motor Company	2.210	19,50%
Volkswagen AG	2.008	17,72%
Stellantis N. V.	1.569	13,84%
Mercedes-Benz Group AG	282	2,49%
Total	**11.334**	**100,00%**

4.2 Generating Topics

4.2.1 Preparation for Latent Dirichlet Allocation

Before running the LDA, the dataset is converted into a corpus and then into a document-feature matrix with the functions of the "quanteda" Package (Benoit et al., 2018). As seen in Table 4, the document-feature matrix consists of the individual patent titles, here called texts in the row labels. Each bigram is represented by a column, here called feature. The matrix structures the word frequencies in a sparse way, to enable efficient processing. The dataset is subset to only include patents filed in 2020 or later, to provide insight into the topics present in the newest data. Later, in the time series, data starting in 2010 to 2023 is used in the LDA.

Table 4

Snippet of the Document-Feature Matrix

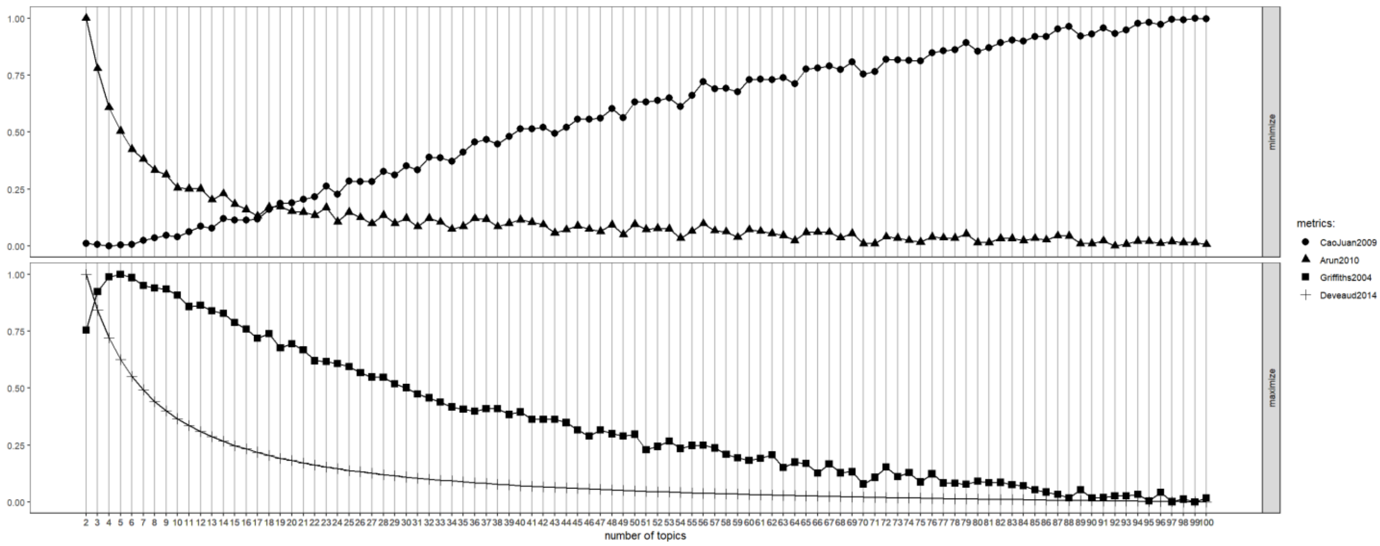
```
       features
docs    e_fire fire_gun gun_fire fire_fight fight_case case_charg charg_object predict_electr
  text2      1        1        1          1          1          1            1              0
  text3      0        0        0          0          0          0            0              1
  text4      0        0        0          0          0          0            0              0
  text5      0        0        0          0          0          0            0              0
  text6      0        0        0          0          0          0            0              0
  text7      0        0        0          0          0          0            0              0
```

4.2.2 Topic Tuning

The main difficulty in implementing the LDA is determining the optimum number of topics, along with trimming the dataset in a way that the LDA produces relevant results. This section discusses this process. The dataset needs to be trimmed to remove the most common features and arrive at a parameter configuration where less frequent and distinguishable topics are still present. The number of topics, which is specified by the user upfront, needs to be determined as well. Since there is no clear indicator of which topic number and trim parameters are the best upfront, the optimal configuration needs to be approximated through trial and error. Each trial is assessed by the researcher regarding topic specificity, coherence of terms within a topic, and overlap across topics. To aid in determining the optimal number of topics, the ‘ldatuning’ package was used (Murzintcev, 2020). This package provides functionality to use four scientific topic number estimation methods. The methods by Arun et al. (2010) and Cao et al. (2009) are topic number minimization methods, whereas the methods by Deveaud et al. (2014) and (Griffiths & Steyvers, 2004) are maximization methods. The four methods were compiled in a graph, which is used in combination with the elbow method (Marutho et al., 2018) to determine the optimal number of topics. The elbow method is a heuristic used to determine the optimal number of topics in a dataset. In the context of this research, the four methods are run on a specified range of topics and plotted by model fit against the number of topics. The ‘elbow’ in the plot indicates the point where adding more topics does not significantly improve the model fit. This point is considered a reasonable choice for the optimal number of topics, as it strikes a balance between model complexity (number of topics) and explanatory power (model fit). As seen in Figure 5, the optimal number of topics is somewhat ambiguous and could lie either at 13 or 17 topics. Less topics are chosen, as more topics led to less interpretability in pre-testing.

Figure 4

Graph of the Four Methods Employed to Determine K



4.2.3 Fitting the LDA model

The “stm” package (Roberts et al., 2019) was used to run the Latent Dirichlet Allocation. Maximum probability, FREX and Score were used to determine the most relevant terms belonging to the topic. The package automatically provides these three metrics. The FREX measure aims to display the terms which are both frequent but exclusive to the topic and may sometimes produce more fitting terms than the probabilistic indicator. Notably, there is strong overlap between the highest probability and FREX. The measure is calculated through the following equation:

Equation 1

FREX Metric by Bischof & Airoldi (2012)

$$\mathrm{FREX} = \left(\frac{w}{F} + \frac{1-w}{E}\right)^{-1}$$

Score is a metric developed by Chang (2015), which is computed according to Equation 3 below.

Equation 2

Score Metric

$$\beta_{v,k}(\log \beta_{w,k} - \frac{1}{K}\sum_{k'} \log \beta_{v,k'})$$

Below, in Table 5, the results of the LDA as an excerpt are displayed. Each topic displays the fifteen top terms according to the three measures presented previously. The earlier a word appears in the list, the more likely it is to belong to the topic. The full output can be found in Appendix 6.1.

Table 5

Results of the LDA. Excerpt of Three Topics

Topic 1 Top Words:

Highest Prob: front_bumper, mount_structur, support_devic, devic_oper, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, devic_includ, measur_devic

FREX: front_bumper, mount_structur, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, speed_control, connect_structur, engin_cool, absorb_member

Score: mount_structur, front_bumper, oper_devic, run_light, daytim_run, support_devic, brake_devic, wheel_suspens, remot_control, devic_oper, instal_structur, hybrid_control, seat_belt, bumper_lower, resist_measur

Topic 2 Top Words:

Highest Prob: exhaust_gas, batteri_cell, display_control, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, travel_control, exhaust_line, instrument_panel, forc_control, devic_exhaust, electr_heat, gas_recircul

FREX: exhaust_gas, batteri_cell, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, exhaust_line, devic_exhaust, electr_heat, gas_recircul, seat_compris, heat_devic, exhaust_aftertreat, electr_batteri

Score: exhaust_gas, gas_purif, purif_devic, display_control, gas_aftertreat, batteri_cell, electr_heat, exhaust_aftertreat, thermal_manag, gas_recircul, heat_catalyst, aftertreat_combust, catalyst_devic, exhaust_line, travel_control

Topic 3 Top Words:

Highest Prob: devic_motor, least_one, solid_state, state_batteri, devic_program, server_devic, autom_drive, cool_structur, map_generat, motor_oper, user_interfac, electr_drive, front_door, support_structur, radar_sensor

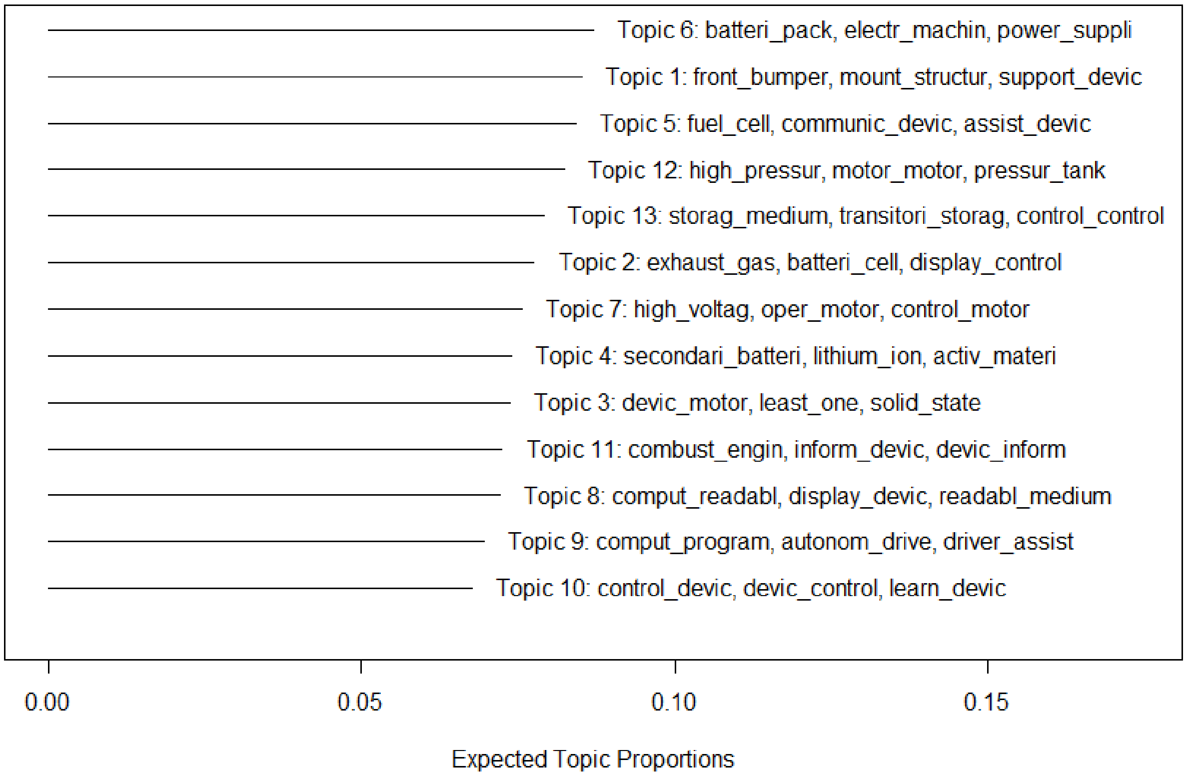
FREX: solid_state, state_batteri, devic_program, autom_drive, cool_structur, map_generat, electr_drive, front_door, support_structur, radar_sensor, sensor_devic, radiat_grill, electr_circuit, interior_trim, cruis_control

Score: devic_motor, state_batteri, solid_state, map_generat, least_one, devic_program, server_devic, user_interfac, cool_structur, front_door, radar_sensor, batteri_deterior, autom_drive, support_structur, solid_electrolyt

The expected topic proportions of the topics generated, which attempt to quantify the likelihood of a topic being present in a random document, are plotted below in Figure 5. All topics seem to be in a similar range of around 6.5% to 9%.

Figure 5

Expected Topic Proportions of the LDA



4.3 Topic Labelling

The LDA produced 13 unlabeled topics, which need to be labelled. This section describes the method through which topics were labeled, and what steps were taken to make the labeling process robust, non-biased and transparent. The topics generated by the LDA contain bigrams. Together, bigrams belonging to the same topic create the meaning of the topic. Thus, the label of the topic can be inferred from the terms in it. Though not all terms may contribute the same amount of meaning to the topic. This may stem from terms not being understandable well, which means expert knowledge is needed, or an apparent misfit with the other terms. To methodically label topics, the decision-making process on which terms contribute to the topic and which contribute less, is documented and reasoned later in this section. This is done to provide transparency of decisions made and justify the assigned labels. Additionally, to gain a deeper understanding of the terms, terms were manually investigated by the researcher to gain a deeper understanding of the terms. This aids interpretation of the terms, resulting in more precise topic labels.

The assigned labels of the topics are displayed in Table 6. Three topics have a mix of terms within, where two to some degree differing themes appear. These are Topic 1, Topic 3 and Topic 5. The topic Computer-based Systems appears twice, in Topics 8 and 13, with Topic 8 being the clearer one. Lastly, Topic 12 was uncoherent and thus uninterpretable.

Table 6

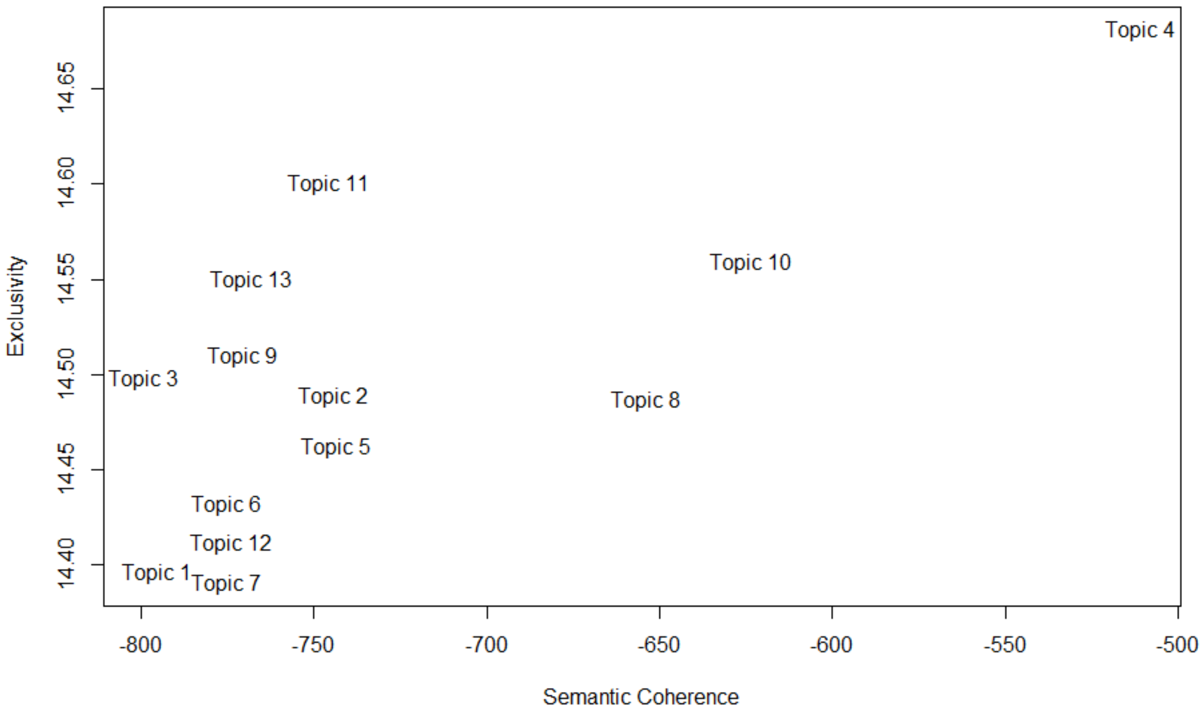
Overview of Assigned Topic Labels

Topic 1	Physical Components and Safety Features
Topic 2	Exhaust Emission Reduction Technology
Topic 3	Battery and Advanced Assistance Systems
Topic 4	Secondary Battery
Topic 5	Fuel Cell and Devices
Topic 6	Electrical Architecture
Topic 7	Powertrain
Topic 8	Computer-based Systems
Topic 9	Advanced Driver Assistance Systems
Topic 10	Control and Diagnostic Devices
Topic 11	Internal Combustion Engine (ICE) and ECU
Topic 12	Uninterpretable
Topic 13	Computer-based Systems (2)

The topics were plotted by exclusivity of words and semantic coherence using the top 15 associated terms in Figure 7. It should be noted that neither axis' starting values are zero. Topic 4 (Secondary Battery) stands out as high scorer across both variables. Notably, Topic 10 and Topic 8 score approximately 100-150 points higher on Semantic Coherence than the remainder of the topics. Regarding Exclusivity, the range of values is very small, suggesting that exclusivity scores are similar across all topics.

Figure 6

Graph of Topics by Exclusivity and Semantic Coherence



In the following section, the topic labels are assigned, by using the keywords to arrive at a clear understanding of the label. Additionally, unfitting keywords are investigated and mentioned for transparency. Lastly, the relevance to and existence in the automotive industry is stated, to tie the topics to reality. For readability, the mentioned terms which are stemmed, are transformed into full words.

Topic 1: Physical Components and Safety Features

Table 7

Top Terms of Topic 1 by Highest Probability, FREX & Score

Topic 1 Top Words:
Highest Prob: front_bumper, mount_structur, support_devic, devic_oper, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, devic_includ, measur_devic
FREX: front_bumper, mount_structur, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, speed_control, connect_structur, engin_cool, absorb_member
Score: mount_structur, front_bumper, oper_devic, run_light, daytim_run, support_devic, brake_devic, wheel_suspens, remot_control, devic_oper, instal_structur, hybrid_control, seat_belt, bumper_lower, resist_measur

The topic focuses on physical components, such as front bumper and mounting structures. Further, it includes brake devices, control systems, seat belts and seemingly daytime light activation, as well as wheel suspension. The label inferred is thus structural elements and safety features. Semantic coherence seems rather low, except for device and control words. The graph in Figure 6 shows Topic 1 to score both second-lowest in semantic coherence and second-worst in exclusivity of terms. The terms are somewhat interpretable though, and commonly mentioned safety features like seat belt create an appearance of a safety theme, also considering brakes, control systems and suspension, as well as bumpers, which are often the first point of contact during an accident and contain a large number of sensors related to safety systems. Safety is an important topic for automotive companies (Mercedes-Benz Group AG, 2022). There is no solid reasoning for the topic label, as it is semantically and exclusivity-wise too diverse.

Topic 2: Exhaust Emission Reduction Technology

Table 8

Top Terms of Topic 2 by Highest Probability, FREX & Score

Topic 2 Top Words:
Highest Prob: exhaust_gas, batteri_cell, display_control, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, travel_control, exhaust_line, instrument_panel, forc_control, devic_exhaust, electr_heat, gas_recircul
FREX: exhaust_gas, batteri_cell, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, exhaust_line, devic_exhaust, electr_heat, gas_recircul, seat_compris, heat_devic, exhaust_aftertreat, electr_batteri
Score: exhaust_gas, gas_purif, purif_devic, display_control, gas_aftertreat, batteri_cell, electr_heat, exhaust_aftertreat, thermal_manag, gas_recircul, heat_catalyst, aftertreat_combust, catalyst_devic, exhaust_line, travel_control

Most terms clearly belong to exhaust emission reduction technologies, as shown by exhaust gas, gas purify, purifying device and gas recirculation. The FREX measure seems to produce the best results, likely because of the exclusivity measure included. Focusing on FREX, 12 of the 15 terms fit the topic clearly. Tracing the bigrams to their original full-text supports this, with titles such as “Exhaust gas purification device, associated exhaust line and purification method”. Most notably, battery-cell does not fit the other terms, and is ranked highly. The full-texts do not offer a direct explanation. One seemingly un-fitting term, electrical heated, is commonly used as shown by this patent title: “Electrically heated catalytic device and method for manufacturing electrically heated catalytic device”. Possibly, the mention of the word electric may have led to the inclusion of other words related to electricity, such as fuel-cell. According to Figure 6, semantic coherence seems to be average, and topic exclusivity too. The terms together create a strong argument through their specificity for the label of this topic. Exhaust emission technologies have become highly relevant for automotive companies, with regulations such as the EURO 6 and coming EURO 7 norms forcing emission reduction (Volkswagen AG, 2023).

Topic 3: Battery and Advanced Assistance Systems

Table 9

Top Terms of Topic 3 by Highest Probability, FREX & Score

Topic 3 Top Words:
Highest Prob: devic_motor, least_one, solid_state, state_batteri, devic_program, server_devic, autom_drive, cool_structur, map_generat, motor_oper, user_interfac, electr_drive, front_door, support_structur, radar_sensor
FREX: solid_state, state_batteri, devic_program, autom_drive, cool_structur, map_generat, electr_drive, front_door, support_structur, radar_sensor, sensor_devic, radiat_grill, electr_circuit, interior_trim, cruis_control
Score: devic_motor, state_batteri, solid_state, map_generat, least_one, devic_program, server_devic, user_interfac, cool_structur, front_door, radar_sensor, batteri_deterior, autom_drive, support_structur, solid_electrolyt

The terms generally have commonalities, in the sense that they are all related to electricity and systems based on it. There are two main themes within this topic, namely batteries for electric cars and advanced assistance systems, such as server device, automatic drive, map generation, user interface, radar sensor, cruise control. Though there are also seemingly non-fitting terms such as front door and radiator grill. The radiator grill usually contains sensors such as for radar and cruise control though. There are terms such as cooling structure, device motor and motor operation, which may not initially seem related, but investigation into the full-text patent titles reveals connections to electric engines, such as these: “Battery cooling structure in vehicle”, “Battery cooling structure in vehicle”. Semantic coherence according to Figure 6 is the lowest, with exclusivity being average. The terms were interpretable, though the topic seems to have two themes. Thus, the label is not fully clear and care should be taken in interpreting findings, especially as there is overlap with Topic 8 (Advanced Driver Assistance Systems) and Topic 4 (Secondary Battery). In the automotive industry, advanced assistance systems have been a focus group of innovation over the last decade, as well as electric vehicles (Stellantis N.V., 2023).

Topic 4: Secondary Battery

Table 10

Top Terms of Topic 4 by Highest Probability, FREX & Score

Topic 4 Top Words:
Highest Prob: secondari_batteri, lithium_ion, activ_materi, storag_devic, batteri_manufactur, ion_secondari, posit_electrod, electrolyt_secondari, power_storag, aqueous_electrolyt, electrod_activ, batteri_batteri, negat_electrod, data_collect, manufactur_electrod
FREX: secondari_batteri, lithium_ion, activ_materi, storag_devic, batteri_manufactur, ion_secondari, posit_electrod, electrolyt_secondari, power_storag, aqueous_electrolyt, electrod_activ, negat_electrod, data_collect, manufactur_electrod, electrod_plate
Score: secondari_batteri, lithium_ion, ion_secondari, posit_electrod, activ_materi, electrolyt_secondari, aqueous_electrolyt, electrod_activ, storag_devic, batteri_manufactur, power_storag, negat_electrod, data_collect, electrod_plate, nonaqu_electrolyt

The terms clearly belong to secondary batteries, as nearly all terms describe parts of it, such as the manufacturing method, or the components it is made of, like lithium-ions, positive and negative electrodes. Only one term does not match the label, namely data collection. According to Figure 6, topic exclusivity and semantic coherence are the highest of all topics. The terms topically fit nearly perfectly to each other, due to them all pertaining to secondary batteries. Resultingly, interpretability is high. Electric cars, and its main component, the battery, have been at the forefront of automotive companies' innovation efforts for years now (Mercedes-Benz Group AG, 2022; Stellantis N.V., 2023; Volkswagen AG, 2023).

Topic 5: Fuel Cell and Devices

Table 11

Top Terms of Topic 5 by Highest Probability, FREX & Score

Topic 5 Top Words:
Highest Prob: fuel_cell, communic_devic, assist_devic, drive_assist, light_devic, inform_provid, devic_seat, cell_devic, steer_devic, fuel_tank, cell_control, cell_stack, control_fuel, cell_fuel, fix_structur
FREX: fuel_cell, communic_devic, drive_assist, light_devic, inform_provid, cell_devic, steer_devic, fuel_tank, cell_control, cell_stack, control_fuel, cell_fuel, park_assist, automat_park, board_electr
Score: fuel_cell, inform_provid, cell_devic, cell_fuel, cell_stack, cell_control, steer_devic, drive_assist, control_fuel, assist_devic, light_devic, devic_seat, valet_park, automat_park, communic_devic

The terms seem to indicate fuel cells and generally devices as themes. The fuel cell theme is supported by fuel-cell being the first term according to each measure, but also related bigrams such as cell device, cell control, cell stack, cell fuel. Then there is the devices theme, which may stem from fuel cell patents often calling it fuel cell device. In turn, other types of devices, such as lighting device, communication device, assistance device may have been consolidated into this topic because of that. Semantic coherence and exclusivity according to Figure 6 are average. In this case, there is underlap between the terms (two themes) and some overlap to other topics, such as Topic 3 with advanced assistance systems being mentioned in this topic too. While it is clear that the topic involves fuel cells and other devices, it is not exclusive to one theme, and thus not clearly interpretable.

Nevertheless, fuel cells as a topic of innovation have been explored by automotive companies for more than a decade now and are still relevant due to the need for sustainability (Toyota Motor Corporation, 2022). Thus, the topic itself is backed by evidence from the automotive industry, notwithstanding the mixed terms in it.

Topic 6: Electrical Architecture

Table 12

Top Terms of Topic 6 by Highest Probability, FREX & Score

Topic 6 Top Words:
Highest Prob: batteri_pack, electr_machin, power_suppli, batteri_modul, air_condit, electr_power, devic_electr, electr_energi, brake_control, energi_storag, automat_drive, control_electr, adjust_devic, suppli_devic, evapor_emiss
FREX: batteri_pack, electr_machin, power_suppli, batteri_modul, air_condit, electr_power, brake_control, energi_storag, automat_drive, adjust_devic, evapor_emiss, emiss_control, rotat_electr, airbag_devic, autonom_mobil
Score: batteri_pack, electr_machin, batteri_modul, power_suppli, electr_power, air_condit, brake_control, rotat_electr, evapor_emiss, electr_energi, emiss_control, modul_batteri, energi_storag, automat_drive, devic_electr

The terms pivot toward an electric direction, with battery pack and battery module seemingly suggesting electric batteries as topic again, but the terms in general are different. They are more focused towards the electric architecture of a vehicle. Electrical machine, power supply, air conditioning, electrical power, device electric, electric energy and brake control describe different electrical systems in the car. From this perspective, battery pack can make sense, as it is the source of electricity for the other systems. The terms evaporate emissions and emission control cannot be explained and seemingly do not complement the existing theme. Semantic coherence and exclusivity according to Figure 6 are below average, yet the results are interpretable. The topic of electrical architecture seems to fit most of the terms, and fits the automotive industry, as especially electric systems and their architecture have become increasingly complex over the years and continue to do so (Askaripoor et al., 2022; Burcicki, 2020).

Topic 7: Powertrain

Table 13

Top Terms of Topic 7 by Highest Probability, FREX & Score

Topic 7 Top Words:
Highest Prob: high_voltag, oper_motor, control_motor, steer_wheel, mobil_devic, power_control, light_guid, devic_compris, voltag_batteri, fuel_pump, lower_bumper, agent_control, batteri_electr, power_manag, wheel_drive FREX: high_voltag, oper_motor, control_motor, power_control, voltag_batteri, lower_bumper, agent_control, batteri_electr, power_manag, wheel_drive, compon_motor, drive_behavior, notif_devic, vibrat_damper, power_generat Score: high_voltag, steer_wheel, control_motor, power_control, voltag_batteri, oper_motor, light_guid, mobil_devic, wheel_drive, sourc_estim, generat_locat, abnorm_behavior, drive_behavior, agent_control, behavior_notif

The terms seem to belong mostly to the powertrain of a vehicle. The terms seem to include electrical vehicles as well, but also ICE vehicles. The theme of the powertrain stems from high voltage, operation motor, control motor, voltage battery, fuel pump, power generation and battery electric all belonging to either electric or ICE engines. Then, control motor, power control, power management, driving behavior may relate to control of the powertrain while driving. The theme is supported by wheel drive, which was found to be four or all-wheel drive after investigating full-text patent titles. There are terms not fitting the theme, such as mobile device, light guide, lower bumper, agent control, which are not explainable. Semantic coherence and exclusivity according to Figure 6 are below average, but the results are interpretable to an extent. The powertrain theme is lexically diverse and thus arguable, especially considering the presence of unfitting terms. This means care should be taken in interpreting findings from this topic. Moving on, powertrains are among the most important systems in a car, which resultingly are under strong innovation efforts by the automotive industry (Aaldering et al., 2019; Zapata & Nieuwenhuis, 2010), which suggests that the theme exists and is important for automotive companies.

Topic 8: Computer-based Systems

Table 14

Top Terms of Topic 8 by Highest Probability, FREX & Score

Topic 8 Top Words:
Highest Prob: comput_readabl, display_devic, readabl_medium, record_medium, transitori_comput, drive_devic, manag_devic, protect_devic, devic_display, readabl_storag, readabl_record, motor_correspond, devic_comput, medium_control, lock_devic FREX: comput_readabl, display_devic, readabl_medium, transitori_comput, drive_devic, manag_devic, protect_devic, readabl_record, cup_holder, motor_display, rear_door, devic_manag, manag_program, pressur_vessel, air_condition Score: comput_readabl, display_devic, readabl_medium, transitori_comput, record_medium, drive_devic, readabl_record, manag_devic, devic_display, readabl_storag, protect_devic, cup_holder, motor_display, ion_batteri, steer_wire

The majority of terms relate to computer-based systems, such as computer readable, display device, readable medium, recording medium, transitory computer. These terms are part of the operating process of a computer in the various forms it is used in a car, such as a display device, a readable storage device or a control device.

There are some terms that do not fit, such as cup holder, which investigation into the full-text patent titles could not explain. Nevertheless, the theme is well interpretable, which is supported by the above average score of semantic coherence and average score of exclusivity, as seen in Figure 6. The theme of computer-based systems is also found in the automotive industry, where computer-based systems are central to innovation efforts regarding electrification, digitalization, autonomous functions and vehicle connectivity (Mercedes-Benz Group AG, 2022; Stellantis N.V., 2023).

Topic 9: Advanced Driver Assistance Systems

Table 15

Top Terms of Topic 9 by Highest Probability, FREX & Score

Topic 9 Top Words:
Highest Prob: comput_program, autonom_drive, driver_assist, control_hybrid, shock_absorb, devic_storag, drive_control, hybrid_electr, detect_devic, oper_manag, power_transmiss, pre_chamber, program_product, devic_comput, monitor_devic
FREX: autonom_drive, driver_assist, control_hybrid, drive_control, hybrid_electr, oper_manag, power_transmiss, pre_chamber, hold_devic, ventil_devic, devic_autonom, lane_chang, self_drive, control_engin, fuel_vapor
Score: comput_program, autonom_drive, driver_assist, drive_control, control_hybrid, shock_absorb, hybrid_electr, oper_manag, pre_chamber, devic_storag, road_obstacl, obstacl_detect, fuel_vapor, self_drive, lane_chang

Many terms relate to advanced driver assistance systems as possible theme, such as autonomous drive, driver assist, self-drive, lane change, road obstacle, obstacle detection. The terms of systems responsible for executing these functions are also present, such as computer program, program product, device storage, detection device. There are some related terms, such as control hybrid, shock absorber and power transmission, which might be controlled by advanced driver assistance systems. Though there are some unfitting terms, which could also not be explained after investigation, such as pre-chamber and fuel vaporization. It is also unclear why the term hybrid appears. Semantic coherence and exclusivity are average, according to Figure 6. Despite these possible issues, the first terms of the topic are clearly related to autonomous driving, allowing for an evidence-based interpretation. Innovations in advanced driver assistance systems are strongly pursued by automotive companies (Mercedes-Benz Group AG, 2022; Stellantis N.V., 2023; Volkswagen AG, 2023) and thus the topic is relevant to automotive companies.

Topic 10: Control and Diagnostic Devices

Table 16

Top Terms of Topic 10 by Highest Probability, FREX & Score

Topic 10 Top Words:
Highest Prob: control_devic, devic_control, learn_devic, estim_devic, data_analysi, devic_power, analysi_devic, abnorm_detect, detect_devic, control_data, damp_control, motor_control, fuel_suppli, heat_exchang, air_fuel
FREX: control_devic, devic_control, learn_devic, control_data, damp_control, air_fuel, control_storag, devic_learn, diagnosi_devic, fuel_ratio, generat_control, control_learn, softwar_updat, devic_hybrid, front_lower
Score: control_devic, devic_control, data_analysi, analysi_devic, control_data, learn_devic, damp_control, fuel_ratio, air_fuel, control_learn, devic_learn, abnorm_detect, generat_control, ratio_sensor, devic_power

The terms seem to fit a theme of devices used to control and diagnose processes. The terms control device, device control, learning device, estimate device, data analysis, analysis device, abnormality detection, detection device, control data, damping control, motor control often contain the word control and words related to diagnostics, such as abnormality detection. These might be used in damping, or the air-fuel ratio as the following patent titles suggest: "DAMPER CONTROL DEVICE FOR VEHICLE, DAMP CONTROL SYSTEM, DAMPER CONTROL METHOD AND DATA PROVISION DEVICE" and "Abnormality detection device for air-fuel ratio sensor, abnormality detection system for air-fuel ratio sensor, data analysis device, and control device for internal combustion engine". Semantic coherence scores second best of all topics, with exclusivity being above average. It seems likely based on the terms, that the label of the topic is related to and interpretable as control and diagnostic devices. In the automotive industry, control and diagnostic devices have risen in importance with the introductions of legal minimum standards and the ever-increasing complexities of cars (Oliveira et al., 2017; Saibannavar et al., 2021), thus making the topic relevant to the automotive industry.

Topic 11: Internal Combustion Engine (ICE) and Engine Control Units

Table 17

Top Terms of Topic 11 by Highest Probability, FREX & Score

Topic 11 Top Words:
Highest Prob: combust_engin, inform_devic, devic_inform, devic_combust, engin_control, object_detect, artifici_muscl, batteri_motor, detect_devic, inform_program, two_stroke, misfir_detect, control_combust, stroke_uniflow, drive_forc
FREX: combust_engin, devic_inform, engin_control, object_detect, artifici_muscl, two_stroke, misfir_detect, control_combust, stroke_uniflow, drive_forc, uniflow_scaveng, engin_oper, center_consol, batteri_case, oper_combust
Score: combust_engin, devic_combust, inform_devic, devic_inform, two_stroke, engin_control, artifici_muscl, misfir_detect, stroke_uniflow, uniflow_scaveng, object_detect, control_combust, oper_combust, larg_turbocharg, inform_program

The theme seems to be about internal combustion engines, with the inclusion of engine control units and other information systems used to optimize engine performance. The terms combustion engine, device combustion, two stroke, misfire detection, control combustion, stroke uniflow and engine operation all relate to ICE, with two stroke engine configurations also mentioned. The presence of two stroke engines seems surprising, given that there are no new production cars with this type of engine for sale in Europe and the U.S. due to not being able to meet environmental regulations (Martini et al., 2009). Nevertheless, Volkswagen AG holds 25 patents related to two-strokes. Additionally, since only the four-stroke engine is used in a modern car, the patent title does not explicitly mention the stroke. The theme is broadened by the inclusion of engine control units, whose label is interpreted through terms such as information device, device information, engine control, detection device, information program, misfire detection and control combustion. Now follows an example patent title: “Misfire detection device for internal combustion engine, misfire detection system for internal combustion engine, data analysis device, and controller for internal combustion engine”. There are terms not fitting the theme, such as object detection, which might stem from its potential semantic similarity to misfire detection technology. Further, artificial muscle is a muscle imitation for robots by Toyota. This highlights that while the patent titles are from automotive companies, some of these are serving other industries too. The discussion section elaborates further on limitations. Semantic coherence and exclusivity in Figure 6 are above average, and the theme is clearly interpretable. While there may be discussions on the degree of internal fit within the topic, it seems obvious that it has to do with ICE and related digital systems. In the automotive industry, the ICE has been the dominant engine type for more than a hundred years (Ojapah et al., 2013). While the end of its lifecycle is in sight and other engine types are gaining market share, the ICE is still highly relevant to automotive companies (Sinigaglia et al., 2022).

Topic 12: Uninterpretable

Table 18

Top Terms of Topic 12 by Highest Probability, FREX & Score

Topic 12 Top Words:
Highest Prob: high_pressur, motor_motor, pressur_tank, automat_transmiss, tank_manufactur, neural_network, devic_communic, transmiss_control, seat_seat, manufactur_high, storag_compart, least_partial, front_grill, batteri_control, cylind_head
FREX: high_pressur, pressur_tank, automat_transmiss, tank_manufactur, transmiss_control, manufactur_high, storag_compart, front_grill, batteri_control, cylind_head, inject_mold, manufactur_devic, support_element, occup_protect, charg_batteri
Score: pressur_tank, high_pressur, automat_transmiss, motor_motor, manufactur_high, tank_manufactur, storag_compart, transmiss_control, neural_network, inject_mold, cylind_head, front_grill, restraint_devic, tank_high, devic_communic

The terms show no major themes, instead there is the theme of high pressure tanks, transmissions, electronic systems such as device communication and neural networks and seemingly random other terms like seat seat, front grille, cylinder head, battery control. According to Figure 6, semantic coherence and exclusivity are below average. Despite these values according to the two measures, the terms are not interpretable enough to derive a label from. Thus, no relevance towards the automotive industry can be drawn and the topic will be excluded from following steps.

Topic 13: Computer-based Systems (2)

Table 19

Top Terms of Topic 13 by Highest Probability, FREX & Score

Topic 13 Top Words:
Highest Prob: storag_medium, transitori_storag, control_control, control_program, inform_inform, medium_store, rear_bumper, electr_motor, traction_batteri, determin_devic, bumper_cover, devic_determin, drive_support, inform_manag, trim_element
FREX: transitori_storag, inform_inform, medium_store, rear_bumper, electr_motor, bumper_cover, drive_support, inform_manag, trim_element, engin_start, balanc_train, store_program, communic_communic, relay_devic, carbon_dioxid
Score: storag_medium, inform_inform, transitori_storag, control_control, medium_store, rear_bumper, control_program, balanc_train, bumper_cover, electr_motor, drive_support, relay_devic, train_control, store_program, traction_batteri

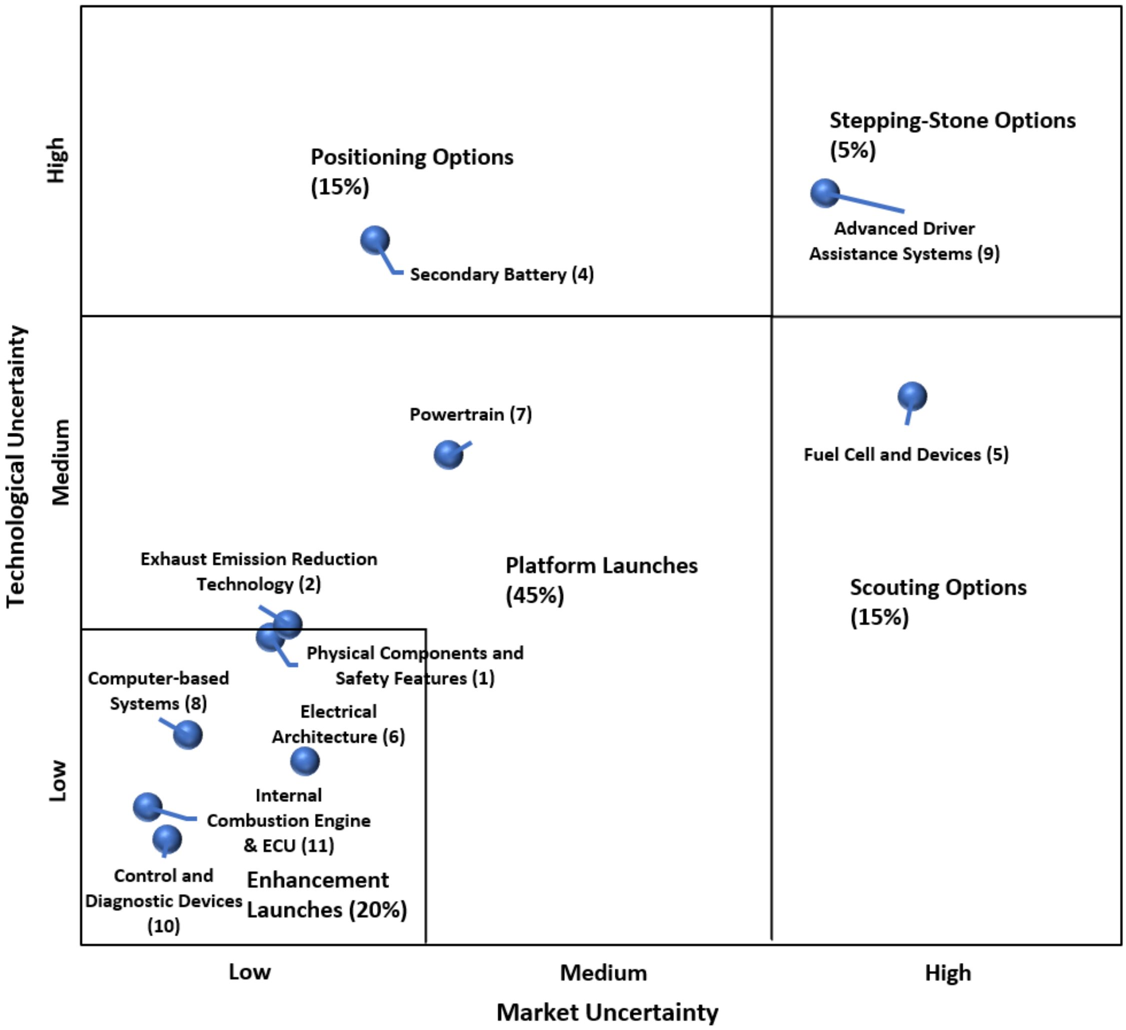
The topic seems to have a theme of computer-based systems similar to Topic 8, but contains a number of irrelevant terms too. Supporting the theme among more are the terms storage medium, transitory storage, control control, control program, information information, medium store. Though a significant number of terms do not fit, such as rear bumper, where only 1 out of 51 patent titles mentions a digital device. Due to the misfit, care should be taken in interpreting the topic, and instead the Topic about Computer-based systems may be taken. According to Figure 6 this topic scores slightly above average on exclusivity and average on semantic coherence. As mentioned in Topic 8, computer-based systems are the basis of innovation in electrification, digitalization, and automation (Mercedes-Benz Group AG, 2022; Stellantis N.V., 2023).

4.4 Applying the R&D Portfolio Matrix on the Topics

The topics are plotted according to their market and technical uncertainty scores in the R&D Portfolio matrix below in Figure 8. Topics 3, 12 & 13 were not included, as Topic 12 is uninterpretable and Topic 13 is a lesser quality repetition of Topic 8. Topic 3 contains two themes, both of which are present separately as well. Thus, Topic 3 does not add value. In general, topics seem to err slightly more on the technological uncertainty side, rather than on market uncertainty. When considering the prescribed percentages of MacMillan & McGrath (2002), the category Enhancement Launches seems to be overrepresented, with Platform Launches being underrepresented followingly. The Positioning Options, Stepping-Stone Options and Scouting Options seem adequately represented, with one topic belonging to each respective category.

Figure 8

Applied R&D Matrix according to MacMillan & McGrath (2002)



Five topics were scored into the enhancement launches category. These topics are Physical Components and Safety Features, Computer-based Systems, Electrical Architecture, Internal Combustion Engine & ECU, and lastly Control and Diagnostic devices. Physical Components and Safety features lay on the tangent of Enhancement- and Platform Launches, which is due to designs of cars changing frequently and safety systems being subject to slight improvements, thus creating low/medium technological uncertainty, while market uncertainty is low. Market uncertainty is mainly influenced by the willingness of customers to pay a premium for safety (Koppel et al., 2008), and by car design attractiveness (Liu et al., 2017). Computer-based Systems are a diverse topic, being used for more simple operations such as controlling a display, to being the basis of advanced driver assistance systems. In general, though, there were many computer-based systems in a car (Walker et al., 2001) more than 20 years ago already, with most of these computers not being state-of-the-art, as the tasks executed do not require such capabilities. While there are exceptions, with new technologies such as autonomy, there majority of computers in a car is subject to low uncertainty due to dominant designs. Electrical Architecture as a topic is of importance for carmakers, as the architecture needs to tie the multitude of computers and other electronic units together in an efficient and effective way (Pelliccione et al., 2017). While this topic is important and many automotive companies such as AB Volvo are investing in renovating and consolidating electrical architecture, it is an explored domain. Thus, it is subject to rather low uncertainty, both market and technologywise. The topic Internal Combustion Engine & Electronical Control Unit (ECU) is subject to low uncertainty due to being widely adopted and technologically mature (Song & Aaldering, 2019). While market uncertainty arises from regulatory hindrances and bans (Danieli et al., 2023), these uncertainties are relatively predictable as bans are announced years before their commencement. Control and Diagnostic Devices are similar in their uncertainties to electrical architecture, being explored and relatively mature technologies, also widely used in other industries. Thus, these devices are very low in uncertainty due to their technological maturity, wide adoption and simplicity compared to other electronic devices performing complex tasks (Isermann, 2001).

In the category of Platform Launches, Powertrain, and Exhaust Emission Reduction Technology are present. Powertrains are subject to medium technological uncertainty, as tighter emission laws and shortened development cycles have pushed companies to improve these heavily. Furthermore, the electrification and further smartification have resulted in advanced calibration techniques and significantly increased complexity (Wang et al., 2023). Market uncertainty is slightly lower, due to not being a key feature to customers, abilities of powertrains being comparable across competitors and any engine configuration requiring a powertrain. Exhaust Emission Reduction Technology scores barely on medium uncertainty and low on market uncertainty. Such technologies aim to reduce emissions, which are the key innovation focus for automotive companies due. The market is relatively certain, as it is closely related to internal combustion engines. Catalytic converters, one of the main exhaust emission reduction technologies, have been mandatory for petrol cars in the European Union since 1993 with the introduction of the EURO 1 standard. Since then, these systems have evolved technologically, playing a key role in staying below emission limits, which means that companies will continue to evolve the technology, thus creating uncertainty for competitors.

In the Positioning Options category, the topic Electric Battery is situated, which is highly complex technologically, and there are multiple types of batteries, such as Lithium-Ion (Zeng et al., 2019). While some battery types are technologically mature enough to be mass commercialized, when compared to the lifespan of and R&D investment into internal combustion engines, electric batteries are still in its infancy. Due to market demand certainly being high in the future, technological uncertainty is high, as companies need to make electrical vehicles attractive and better than their competitors. Market uncertainty is between medium and low, as this technology is seen as solution to sustainability and climate change, by replacing fossil fuel powered vehicles. This results in market demand being given, as it is supported by governments around the world. Nevertheless, there is some market uncertainty as new charging infrastructure needs to be built, and especially when considering less wealthy countries, infrastructure and the resulting adoption are uncertain (Dioha et al., 2022).

Fuel Cells and related devices are plotted in the Scouting Options Category. Market uncertainty is high, as infrastructure for refilling or replacing fuel cells is not present in most countries. It competes with electrical batteries as solution to climate change but needs different infrastructure. Considering that electric batteries are being adopted at a higher rate, as automotive companies mostly offer electric vehicles rather than fuel-cell powered ones (Van De Kaa et al., 2017), this sways focus towards electric battery development. Nevertheless, there are fuel-cell vehicles on the market, meaning the market is not completely unknown. Technologically, there is medium uncertainty, as the technology has been successfully implemented. But due to the competing engine types, it needs to outperform them in some ways to become a dominant or at least commonly adopted technology (Asif & Schmidt, 2021). Due to the technology not meeting the viability threshold, technological uncertainty is created, as new ways need to be discovered which enhance the technology and boost it to a competitive level.

Lastly, the Stepping-Stone Options category contains the topic Advanced Driver Assistance Systems, which scores high on both technological and market uncertainty. Technological uncertainty is caused by the novelty of artificial intelligence and complexity of it. The novelty of artificial intelligence and its seemingly endless capabilities given the knowledge how to make it work create new pathways for innovation in many aspects of a car. Such possibilities include autonomous functions, such as fully self-driving cars, but also enhanced crash prediction systems, vehicle control systems and lower emissions through better route selection (Abduljabbar et al., 2019). Additionally, such advanced driver assistance systems require highly complex technical configurations and high reliability to be successfully implemented. Thus, the hardware aspect of incorporating these systems leads to additional technological uncertainty, as this adds another layer of complexity to integrate seamlessly into the cars existing systems. Market uncertainty is high, as it is not known what the trajectory of these advanced systems is, and how artificial intelligence may develop. Due to its inestimable capabilities, it is not clear where the market might develop, and customers are not wholly decided yet on the attractiveness of autonomous cars (Meidute-Kavaliauskiene et al., 2021).

4.5 Analysis of Company-Specific Topic-Patent Distribution

Table 20

Topic-Patent Distribution of Toyota Motor Corporation

Topic	Relative Frequency	Frequency
Topic 4 Secondary Battery	12,09%	464
Topic 6 Electrical Architecture	11,96%	459
Topic 5 Fuel Cell and Devices	11,70%	449
Topic 1 Physical Components and Safety Features	10,94%	420
Topic 11 Internal Combustion Engine (ICE) and ECU	9,30%	357
Topic 2 Exhaust Emission Reduction Technology	9,15%	351
Topic 8 Computer-based Systems	8,94%	343
Topic 9 Advanced Driver Assistance Systems	8,86%	340
Topic 7 Powertrain	8,62%	331
Topic 10 Control and Diagnostic Devices	8,44%	324
Total	**100%**	**3838**

Toyota Motor Corporation seems to have somewhat of a balance of frequencies across all topics, with slight, as seen in Table 20. There appears to be an emphasis on electricity through the first two topics, but also a general focus on new engine types, namely batteries and fuel cells. Taking into account the R&D Matrix, two of the three high uncertainty topics are present in the focus group of Toyota.

Table 21

Topic-Patent Distribution of Ford Motor Company

Topics	Relative Frequency	Frequency
Topic 6 Electrical Architecture	15,73%	252
Topic 1 Physical Components and Safety Features	13,48%	216
Topic 7 Powertrain	13,17%	211
Topic 5 Fuel Cell and Devices	12,80%	205
Topic 2 Exhaust Emission Reduction Technology	9,99%	160
Topic 9 Advanced Driver Assistance Systems	7,68%	123
Topic 8 Computer-based Systems	7,62%	122
Topic 4 Secondary Battery	7,43%	119
Topic 11 Internal Combustion Engine (ICE) and ECU	6,37%	102
Topic 10 Control and Diagnostic Devices	5,74%	92
Total	**100%**	**1602**

Ford Motor Company's topic distribution shows dominating and less prominent topics, as seen in Table 21. The electrical architecture of vehicles is the most common topic for Ford, with physical components and safety features, as well as the powertrain and fuel cells following closely behind. Notably, secondary batteries, and advanced driver assistance systems are less focused by Ford, though they score high on uncertainty in the R&D matrix. Additionally, internal combustion engines are also at the bottom of the distribution.

Table 22

Topic-Patent Distribution of Volkswagen AG

Topics	Relative Frequency	Frequency
Topic 1 Physical Components and Safety Features	12,75%	194
Topic 5 Fuel Cell and Devices	11,77%	179
Topic 11 Internal Combustion Engine (ICE) and ECU	11,64%	177
Topic 6 Electrical Architecture	11,31%	172
Topic 8 Computer-based Systems	10,72%	163
Topic 2 Exhaust Emission Reduction Technology	10,59%	161
Topic 7 Powertrain	10,26%	156
Topic 9 Advanced Driver Assistance Systems	7,96%	121
Topic 4 Secondary Battery	7,03%	107
Topic 10 Control and Diagnostic Devices	5,98%	91
Total	**100%**	**1521**

Volkswagen AG according to Table 22 exhibits a largely balanced distribution, with a drop in the last three topics. The company's most frequent patent topic are physical components and safety features, fuel cells and internal combustion engines, though the first seven topics all have a similar share. The last three topics provide a contrast, as advanced driver assistance systems and secondary batteries, either stepping-stone or positioning options seem to be less common, as well as control and diagnostic devices. Though fuel cells, a scouting option, are among the top three topics.

Table 23

Topic-Patent Distribution of Stellantis N.V.

Topics	Relative Frequency	Frequency
Topic 1 Physical Components and Safety Features	16,10%	180
Topic 2 Exhaust Emission Reduction Technology	13,51%	151
Topic 6 Electrical Architecture	12,08%	135
Topic 7 Powertrain	11,09%	124
Topic 5 Fuel Cell and Devices	10,64%	119
Topic 8 Computer-based Systems	8,32%	93
Topic 11 Internal Combustion Engine (ICE) and ECU	7,87%	88
Topic 4 Secondary Battery	7,33%	82
Topic 9 Advanced Driver Assistance Systems	7,25%	81
Topic 10 Control and Diagnostic Devices	5,81%	65
Total	**100%**	**1118**

The topic-patent distribution of Stellantis N.V. as seen in Table 23 shows differences in topic frequencies. Physical components and safety features stand out as most common topic, with exhaust emission reduction technologies following behind. Electrical architecture, the powertrain and fuel cells seem to be similar to each other in share. On the lower end of the distribution, two engine types, namely the ICE and electric engine are found and seemingly less prominent. Advanced driver assistance systems do also not seem to be of major proportion.

Table 24

Topic-Patent Distribution of Mercedes-Benz Group AG

Topics	Relative Frequency	Frequency
Topic 5 Fuel Cell and Devices	19,19%	38
Topic 6 Electrical Architecture	14,14%	28
Topic 7 Powertrain	11,11%	22
Topic 1 Physical Components and Safety Features	10,61%	21
Topic 4 Secondary Battery	10,10%	20
Topic 11 Internal Combustion Engine (ICE) and ECU	8,59%	17
Topic 2 Exhaust Emission Reduction Technology	8,08%	16
Topic 10 Control and Diagnostic Devices	7,07%	14
Topic 9 Advanced Driver Assistance Systems	7,07%	14
Topic 8 Computer-based Systems	4,04%	8
Total	**100%**	**198**

The topic-patent distribution of Mercedes-Benz Group AG is shown in Table 24, and needs to be carefully interpreted, as the total number of patents per topic is low and thus less reliable on for comparisons (Button et al., 2013). Fuel cells seem to be the dominant topic for Mercedes-Benz, with electrical architecture also being slightly more prominent than the following topics. In general, the topic distribution is more uneven when considering the high and low scores, but there is a relatively smooth middle range present.

4.6 Company Strength per Topic-Share

The following section analyzes the relative frequency of each topic in each company's own topic distribution compared to the other companies. This means that the shares are taken from the company's relative frequency, not absolute number, as Toyota Motor Corporation would dominate each topic, and the comparison would be futile. This is done to enable a level comparison across companies, despite differences in total patent numbers.

Table 25

Relative Topic Frequencies of Topic 1 & 2

	Topic 1 Physical Components and Safety Features	Topic 2 Exhaust Emission Reduction Technology
Toyota Motor Corporation	10,94%	9,15%
Ford Motor Company	13,48%	9,99%
Volkswagen AG	12,75%	10,59%
Stellantis N. V.	16,10%	13,51%
Mercedes-Benz Group AG	10,61%	8,08%

In Topic 1 as seen in Table 25, the company with the highest portfolio percentage is Stellantis N.V.. Ford Motor Company and Volkswagen AG are on a similar level ranking second and third, with Toyota Motor Corporation and Mercedes-Benz Group AG ranking similarly too in fourth and fifth place.

Topic 2 in the same table concerns exhaust emission reduction technology and Stellantis N.V. ranks highest again. The other companies are all together in a similar range.

Table 26

Relative Topic Frequencies of Topic 4 & 5

	Topic 4 Secondary Battery	Topic 5 Fuel Cell and Devices
Toyota Motor Corporation	12,09%	11,70%
Ford Motor Company	7,43%	12,80%
Volkswagen AG	7,03%	11,77%
Stellantis N. V.	7,33%	10,64%
Mercedes-Benz Group AG	10,10%	19,19%

In Table 26, Topic 4 concerning the secondary battery is led by Toyota Motor Corporation, with Mercedes-Benz Group AG also having a significantly larger share than the remaining three companies, which score similarly.

In the same table, Topic 5 shows Mercedes-Benz Group AG to be the clear leader in fuel cells and devices. The other companies are distributed close together in the 10-12% range.

Table 27

Relative Topic Frequencies of Topic 6 & 7

	Topic 6 Electrical Architecture	Topic 7 Powertrain
Toyota Motor Corporation	11,96%	8,62%
Ford Motor Company	15,73%	13,17%
Volkswagen AG	11,31%	10,26%
Stellantis N. V.	12,08%	11,09%
Mercedes-Benz Group AG	14,14%	11,11%

Table 27 begins with Topic 6 about the electrical architecture of a vehicle. Ford Motor Company scores highest, with Mercedes-Benz Group AG behind. The remaining companies all score similarly to each other, and a little lower than Mercedes-Benz Group AG.

Topic 7 concerns the powertrain, in which Ford Motor Company leads again, with Volkswagen AG, Stellantis N.V. and Mercedes-Benz Group AG scoring similarly. Notably, Toyota Motor Corporation has the lowest distribution share by a margin.

Table 28

Relative Topic Frequencies of Topic 8 & 9

	Topic 8 Computer-based Systems	Topic 9 Advanced Driver Assistance Systems
Toyota Motor Corporation	8,94%	8,86%
Ford Motor Company	7,62%	7,68%
Volkswagen AG	10,72%	7,96%
Stellantis N. V.	8,32%	7,25%
Mercedes-Benz Group AG	4,04%	7,07%

Table 28 depicts computer-based systems in Topic 8, where Volkswagen AG holds the largest share. All other companies except Mercedes-Benz Group AG score similarly, with this company scoring well below all others.

Topic 9 in the same table is about advanced driver assistance systems, which exhibit an even distribution across all companies with only slight differences.

Table 29

Relative Topic Frequencies of Topic 10 & 11

	Topic 10 Control and Diagnostic Devices	Topic 11 Internal Combustion Engine (ICE) and Engine Control Units
Toyota Motor Corporation	8,44%	9,30%
Ford Motor Company	5,74%	6,37%
Volkswagen AG	5,98%	11,64%
Stellantis N. V.	5,81%	7,87%
Mercedes-Benz Group AG	7,07%	8,59%

Topic 10 in Table 29 shows Toyota Motor Corporation to be the leader in control and diagnostic devices according to the portfolio shares. The other companies display lower shares, with Mercedes-Benz Group AG scoring slightly higher than the others.

The last topic, Topic 11 concerning the internal combustion engine and engine control units shows Volkswagen AG to be the clear leader, with Toyota Motor Corporation following behind. Ford Motor Company scores the lowest, somewhat lagging behind Stellantis N.V..

5. Discussion

Through Latent Dirichlet Allocation, ten viable topics were discovered in the patents. These topics were labelled and inducted into the R&D Portfolio matrix by MacMillan & McGrath (2002). The resulting matrix shows that the topics somewhat resemble the prescribed percentages per category by the authors, with some deviations. The individual company's topic-patent distributions and cross-company comparisons highlight focused and less focused topics. The research questions can now be answered. The first question was "*How can unsupervised text mining uncover current technology themes in the automotive industry?"*, which was demonstrated through using topic modeling uncovering the current technology themes in the automotive industry. Furthermore, the related second research question "*What are the current technologies in the automotive industry?"* was answered as well. The current technologies in the automotive industry were shown in Table 6 in the Topic Labelling section (4.3). The third research question "*How do the individual portfolio compositions of the leading companies vary?"* was answered in section 4.5 and 4.6. It became clear that there are differences between the companies regarding the way they distribute their portfolio, such as by focusing technologies.

The relevance of portfolio management in the context of a global organization should not be underestimated, especially in VUCA environments (Eckert & Hüsig, 2021). Firstly, portfolio management needs to be aligned with the organizations' strategy. By selecting projects aligned with its long-term goals, portfolio management ensures that R&D investments contribute to the success and competitiveness of the firm (Iamratanakul et al., 2009). Moreover, risks are better mitigated through the diversification of risk across the project portfolio (Paquin et al., 2015). Through balancing highly uncertain and more certain R&D projects, this is achieved. In addition, using project portfolio techniques such as the uncertainty matrix used in this research, helps visualize the portfolio to evaluate whether it is optimally diversified. Given more turbulent environments that companies operate in, portfolio management can be used to make innovation more flexible, to be more adapted to changing market conditions and technological disruptions. This can be achieved through approaches such as real options, which is again a use-case of the uncertainty matrix used (Sirmon et al., 2007).

Another important aspect is the ability to evaluate the performance of R&D projects and the whole portfolio through using portfolio management (Eckert & Hüsig, 2021). Evaluating performance helps to identify successful projects, understand reasons for failures, and generally also contributes to the organizational learning process. Through measuring performance, decision-making can be improved, as portfolio management provides a holistic perspective of the R&D portfolio. This is due to being able to assess trade-offs, the current state of the R&D portfolio, and its alignment to company strategy. Lastly, the limited resources available can be optimally allocated to R&D projects through portfolio management, helping to maximize the value and impact of R&D (Sirmon et al., 2007).

5.1 Categories in the Uncertainty Matrix

The theoretical implications of the seemingly skewed portfolio distribution are discussed in detail in this section. Whether these theoretical implications apply to the automotive industry, and what the reasons for a potential misfit are, is discussed in section 5.2. Looking at the results, the constructed R&D portfolio matrix is too dense in the Enhancement Launch category according to theory, which suggests that the analyzed automotive companies are focusing too much on incremental, low-uncertainty launches. According to MacMillan & McGrath (2002), this poses the risk of missing out on the innovations of the future, leading to lowered competitiveness. As this is an analysis of the five largest automotive companies, this may suggest that less risky innovations are preferred, to sustain their positions in a mature market with high barriers to entry. Furthermore, the topics in this category are mostly well-known technologies, whose difficulty lies innovating the integration of these technologies in a cost-efficient, standardized and reliable way. As cars have become extremely complex machines comprised of hundreds of subsystems, the innovation portfolio is adjusted to consolidate and integrate, which is less common in other industries. This is a characteristic of the automotive industry, which should be considered when creating the optimal innovation portfolio according to the R&D portfolio matrix.

Following the overpopulation of the Enhancement Launches category, the Platform Launches category is strongly underrepresented with two topics only. This may signal that companies are not bold enough to innovate with medium uncertainty. It could also suggest that the prescribed percentage of 45% of all innovations to be a platform launch is too high for this industry, as economies of scale and the mature market make it difficult to find enough viable medium-uncertainty innovation projects to pursue.

In the Positioning Options category, the topic of secondary batteries which power electrical vehicles fits the prescribed percentages by MacMillan & McGrath (2002). This suggests that the companies focusing on electric vehicles have a more optimal innovation portfolio. Through the presence of positioning options, automotive companies can continue to serve the same market but with a different technology. In this case, replacing ICE's with electric batteries will alleviate the problems in adhering to emission laws, and possibly increase the sustainability of cars, while essentially only changing the power source of the vehicle. When considering the PESTEL management framework, there are political, social, environmental and legal demands for automotive companies to become more sustainable, which a positioning option like electric power sources may alleviate.

The industry seems to be congruent to the prescribed percentages of the R&D project portfolio matrix in the Scouting Options category. This category is similarly important to optimizing innovation portfolios, as it provides ways to explore new markets or segments. In this case, market segments could be targeted with fuel cells as an alternative to ICE and electric vehicles.

Stepping-Stone Options are fitting the proportion proposals of the theory by MacMillan & McGrath (2002). The stepping-stone in this case is Advanced Driver Assistance Systems which provide automotive companies with competitive advantages, as the future is thought to become more autonomous and smart. Such technologies may become strong selling points and due to their high uncertainty provide opportunities for automotive companies to innovate within a novel technological domain. This contrasts the usual domains these companies innovate in, as they are largely not new but rather well-known.

5.2 Characteristics of the Automotive Industry

The goal of this research was to uncover the current technologies in the automotive industry, classify them and plot them on the uncertainty matrix. The following paragraphs aim to relate the interpretation of the topic distribution on the R&D matrix to the automotive industry's landscape to provide a more comprehensive understanding and validate the results.

Holistically assessing the R&D Portfolio matrix, it becomes apparent that topics generally lean towards technological uncertainty, rather than market uncertainty. This may be explained through a combination of company-specific and external factors. Firstly, automotive companies are often risk-averse, as product cycles are long and consume lots of capital. Through the unpredictability that uncertainty introduces, companies tend to make the majority of their innovation efforts Enhancement Launch focused, while pursuing riskier options on the side. Furthermore, as mentioned before, the maturity of the market also signifies that customer preferences are known and change thus less radical. From another perspective, if the largest automotive companies all choose to not pursue radicality and rather keep market share, it enables each individual company to be complacent, as no disruptors are threatening the industry.

Keeping a steady stream of revenue is thus preferred to having an unpredictable stream of revenue, which due to long product cycles could become a significant financial problem if the innovation is not perceived as attractive by customers. From a customer's view, a car must fulfill their needs, which are most frequently reliability and safety. Additionally, some value brand status, performance, and a large feature selection. Thus, innovations need to fulfill at least the current expectations, which limits automobiles in the way that the car must fulfill the existing expectations first. Then, innovations can build on top of this. New technologies often start with a worse total offering than existing mature technologies, which is a disadvantage to the average automobile customer, who wants the most capable product. Thus, adoption propensity will be low, as shown by the introduction of electrical vehicles into the market, where customer adoption was extremely low at first due to the higher cost of the vehicle and much shorter range primarily (Rogers et al., 2014) . Customers value incremental innovations, as they build upon features the customer is already familiar with. This enables the customer to better understand what the car offers and thus evaluate it correctly, which means the customer is more likely to buy it (Mugge & Dahl, 2013).

Further, regulations in the automotive industry are strict for some technologies, and generally concerned with safety too. Especially the smartification of cars and introduction of autonomous systems are subject to scrutiny, which is considered by automotive companies. The safety aspect means that unreliable technologies, such as autonomous driving now are not legal to use. This hampers the success of new technologies and ability to improve upon these technologies through implementation on the road. Though in some countries, these hindrances do not exist. Furthermore, sustainability is a problem in innovation for automotive companies, as it introduces another requirement that needs to be fulfilled. Generally, there are requirements innovations need to fulfill, which a radical technology might have difficulties with, due to being in the infancy of its technological maturity cycle (Rogers et al., 2014).

Lastly, the automotive industry is dependent on economies of scale (Wynn-Williams, 2009). Incremental innovations are more easily integrated into the manufacturing process, which allows companies to use their existing infrastructure and thus keep cost efficiency high. Building new infrastructure would require large investments and require the radical innovation to be sold frequently, as economies of scale depend on a large number of sales. Though radical modular innovations might mitigate these problems, which would again introduce another requirement on the innovation.

Concluding the section on how the automotive industry's characteristics might explain the results from the R&D portfolio matrix, it has become clear that there are a number of factors influencing automotive companies in the way that they structure their innovation portfolio. The market is saturated, and customers have clear expectations and are less welcoming to new technologies, there are legal and environmental pressures and regulations to follow, and incremental innovations have more synergy with the necessary economies of scale and provide more predictable streams of revenue. These characteristics seem to fit the results of the R&D matrix, with most projects being low in uncertainty, and only a select few being highly uncertain.

5.3 Portfolio Compositions

There are differences between the five companies regarding the topics they focus on in their patents. These differences are especially interesting to observe in the medium to high uncertainty topics, as theory suggests that these are the drivers of competitive advantage and success of the future.

The results show that Toyota Motor Corporation leads the high uncertainty topics of secondary batteries and advanced driver assistance systems, while being average in fuel cells. They also focus less on physical components and the powertrain, while leading the control and diagnostic devices category. This suggests that Toyota places the most importance of the compared companies on high uncertainty options. Additionally, the company holds the largest number of patents in each category, which suggests that they innovate more than its competitors. These arguments would suggest that Toyota Motor Corporation will produce the best innovations in the future.

Ford Motor Company stands out in the comparison in a few topics, scoring arguably first in fuel cells which is elaborated on later. Further, it seems to be strong in powertrain innovation, and electrical architecture, which are two highly important innovation topics that automotive companies face. Especially electrical architecture consolidation is a weakness of many automotive companies, such as Volkswagen AG according to S&P Global (Dixon, 2023). This seems to provide Ford with competitive advantage in these topics. Notably, there are few innovations in ICE, which can be interpreted as an attempt to shift focus to more sustainable power sources, which are highly relevant to future competitiveness (Lukin et al., 2022).

Volkswagen AG stands out in only two topics, namely computer-based systems and ICE. While it has a percentage of its portfolio allocated to each topic, there seems to be more of an even distribution and less focus. It also scores low on electric batteries. While the company is the largest automotive manufacturer in the world, the portfolio distribution raises questions on its future competitiveness. This is especially outlined by the focus on ICE, which will not be a future source of competitive advantage due to regulations and the apparent lack of focus on electric vehicles (European Environment Agency, 2016).

Stellantis N.V. has a relatively balanced portfolio, while standing out in physical components and safety as well as exhaust emission reduction technologies. It scores averagely on new engine types. From this can be interpreted that Stellantis has a balanced portfolio, which does not stand out in high uncertainty topics, but is not necessarily bad.

Mercedes-Benz Group AG holds such a low number of patents, that the results of the comparative analysis are statistically unreliable. Nevertheless, the company scores the best by far in fuel cells. Arguably, the score might be inflated through the bias introduced by the low number of patents. Despite being one of the largest automotive companies, its patents do not reflect the investments into R&D.

5.4 Theoretical and Practical Implications

This research used text mining to gain insights from data relevant for technology foresight, which adds to the existing evidence in academic literature that text mining is a useful approach for R&D, and more specifically also technology foresight (Kayser et al., 2014; Kayser & Blind, 2017b; Mühlroth & Grottke, 2018). Secondly, it fills the research gap of using text mining to create the uncertainty matrix from patent data. This research resultingly contributes to the domain of R&D management and technology foresight by extending the understanding of how text mining can contribute to these two domains. Additionally, since there is a lack of literature concerning the uncertainty matrix, with only one paper by Luo (2011) on the framework, this research extends the understanding of the framework by discussing whether the prescribed allocation percentages of the matrix actually reflect the analyzed industry. Luo (2011) stochastically describes market and technological uncertainty, which are the variables used in the uncertainty matrix, but does not demonstrate how the uncertainty matrix may be constructed from data which this research additionally does.

Regarding business implications and relevance for practitioners, this paper demonstrates that text mining can be used to identify current technology themes from patent data, which might be a viable method in practice to use for detection of technology themes and comparisons of companies' innovation portfolios. This method can be viable as it possesses the capabilities to process large amounts of data in a short time with a relatively light resource expenditure compared to traditional methods such as manually reviewing patents. Due to being unsupervised, it may generate novel topics, which the practitioner is unaware of. It is scalable and can be used on huge datasets, with the flexibility to adjust the algorithm according to the objective, making it a versatile tool. Early judgements on the diversification of the innovation portfolio of a company and industry are made possible through the usage of the uncertainty matrix, which could be valuable information for a practitioner assessing whether there is a need to readjust the portfolio. Lastly, as argued in the discussion, managers should not follow the prescribed percentages too closely, as these percentages might vary between industries. Resultingly, managers should construct a realistic distribution based on their knowledge of the industry or company.

Policymakers may derive value from this research through using the methods themselves to track technologies, which with regards to issues such as climate change can serve as an indicator which technologies the innovation focus seems to be on. Armed with this knowledge, policymakers can roadmap the automotive industry's progress and adjust legislation accordingly.

5.5 Limitations

The low number of patents of Mercedes-Benz Group AG pose the question of how well patents reflect the innovation portfolio, which leads to the limitations of this research. While patents seem to be a viable indicator of innovation (Archibugi, 1992), it does not necessarily mean that this is the case for every company, as exemplified by Mercedes-Benz Group AG. Additionally, this research included only granted patents, which often lag behind, in the way that they are not immediately granted. Further, the datasets are imbalanced regarding the patent share of each company, which makes comparisons only possible on the relative level, as the absolute frequencies would mean that Toyota dominates every topic. Another important fact is that the patents are all from automotive companies, but Toyota for example has patents mentioning artificial muscles in Topic 11 for example, which relates to a robot they built, not an automobile. Thus, the patents do not necessarily belong to the automobile industry, though this is the case for the vast majority.

Latent Dirichlet Allocation may produce better results on longer documents than the short titles used here, but full patents with their description were not available. Additionally, hierarchical clustering was attempted, but not possible due to hardware limitations. Lastly, the labelling was done by the researcher alone, which may have introduced bias. Also, the topics were not validated with automotive experts.

5.6 Recommendations

For future research, the LDA might be compared to multiple different techniques, to find out which methods produce the most relevant results and thus higher quality inputs into the uncertainty matrix. Furthermore, since the automotive industry does not match the theoretical stipulations of the theoretical framework, researchers can investigate whether there really are optimal portfolio distributions. Followingly, it might be researched which specific industry characteristics influence the optimal distribution in which way.

To expand upon the results of this research, managers can pursue a number of options. A time series could be applied to gauge trending topics and use them to foresight change. This would add practical relevance by introducing a timeline, compared to the single point-in-time approach used here. Next, managers aiming to employ the theoretical framework and assess their portfolio based on the theoretical stipulations made by the authors could investigate whether the prescribed portfolio category allocation percentages are realistic for their industry and thus individual circumstances. Additionally, the portfolio uncertainty matrix might be combined with a framework assessing the potential of an innovation, to produce more comprehensive implications for the optimal R&D portfolio. The reasoning is that a holistic assessment of a project portfolio should not be based on the uncertainty of projects alone, but also on their potential, to select the best innovations to pursue.

6. Appendix

6.1 Full LDA Output

Table 29

LDA Output. First Part (Topics 1-6)

Topic 1 Top Words:
Highest Prob: front_bumper, mount_structur, support_devic, devic_oper, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, devic_includ, measur_devic FREX: front_bumper, mount_structur, oper_devic, remot_control, brake_devic, hybrid_control, instal_structur, seat_belt, daytim_run, run_light, wheel_suspens, speed_control, connect_structur, engin_cool, absorb_member Score: mount_structur, front_bumper, oper_devic, run_light, daytim_run, support_devic, brake_devic, wheel_suspens, remot_control, devic_oper, instal_structur, hybrid_control, seat_belt, bumper_lower, resist_measur
Topic 2 Top Words:
Highest Prob: exhaust_gas, batteri_cell, display_control, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, travel_control, exhaust_line, instrument_panel, forc_control, devic_exhaust, electr_heat, gas_recircul FREX: exhaust_gas, batteri_cell, gas_purif, purif_devic, thermal_manag, gas_aftertreat, machin_learn, exhaust_line, devic_exhaust, electr_heat, gas_recircul, seat_compris, heat_devic, exhaust_aftertreat, electr_batteri Score: exhaust_gas, gas_purif, purif_devic, display_control, gas_aftertreat, batteri_cell, electr_heat, exhaust_aftertreat, thermal_manag, gas_recircul, heat_catalyst, aftertreat_combust, catalyst_devic, exhaust_line, travel_control
Topic 3 Top Words:
Highest Prob: devic_motor, least_one, solid_state, state_batteri, devic_program, server_devic, autom_drive, cool_structur, map_generat, motor_oper, user_interfac, electr_drive, front_door, support_structur, radar_sensor FREX: solid_state, state_batteri, devic_program, autom_drive, cool_structur, map_generat, electr_drive, front_door, support_structur, radar_sensor, sensor_devic, radiat_grill, electr_circuit, interior_trim, cruis_control Score: devic_motor, state_batteri, solid_state, map_generat, least_one, devic_program, server_devic, user_interfac, cool_structur, front_door, radar_sensor, batteri_deterior, autom_drive, support_structur, solid_electrolyt
Topic 4 Top Words:
Highest Prob: secondari_batteri, lithium_ion, activ_materi, storag_devic, batteri_manufactur, ion_secondari, posit_electrod, electrolyt_secondari, power_storag, aqueous_electrolyt, electrod_activ, batteri_batteri, negat_electrod, data_collect, manufactur_electrod FREX: secondari_batteri, lithium_ion, activ_materi, storag_devic, batteri_manufactur, ion_secondari, posit_electrod, electrolyt_secondari, power_storag, aqueous_electrolyt, electrod_activ, negat_electrod, data_collect, manufactur_electrod, electrod_plate Score: secondari_batteri, lithium_ion, ion_secondari, posit_electrod, activ_materi, electrolyt_secondari, aqueous_electrolyt, electrod_activ, storag_devic, batteri_manufactur, power_storag, negat_electrod, data_collect, electrod_plate, nonaqu_electrolyt
Topic 5 Top Words:
Highest Prob: fuel_cell, communic_devic, assist_devic, drive_assist, light_devic, inform_provid, devic_seat, cell_devic, steer_devic, fuel_tank, cell_control, cell_stack, control_fuel, cell_fuel, fix_structur FREX: fuel_cell, communic_devic, drive_assist, light_devic, inform_provid, cell_devic, steer_devic, fuel_tank, cell_control, cell_stack, control_fuel, cell_fuel, park_assist, automat_park, board_electr Score: fuel_cell, inform_provid, cell_devic, cell_fuel, cell_stack, cell_control, steer_devic, drive_assist, control_fuel, assist_devic, light_devic, devic_seat, valet_park, automat_park, communic_devic
Topic 6 Top Words:
Highest Prob: batteri_pack, electr_machin, power_suppli, batteri_modul, air_condit, electr_power, devic_electr, electr_energi, brake_control, energi_storag, automat_drive, control_electr, adjust_devic, suppli_devic, evapor_emiss FREX: batteri_pack, electr_machin, power_suppli, batteri_modul, air_condit, electr_power, brake_control, energi_storag, automat_drive, adjust_devic, evapor_emiss, emiss_control, rotat_electr, airbag_devic, autonom_mobil Score: batteri_pack, electr_machin, batteri_modul, power_suppli, electr_power, air_condit, brake_control, rotat_electr, evapor_emiss, electr_energi, emiss_control, modul_batteri, energi_storag, automat_drive, devic_electr

Table 30

LDA Output. Second Part (Topics 7-13)

Topic 7 Top Words:
Highest Prob: high_voltag, oper_motor, control_motor, steer_wheel, mobil_devic, power_control, light_guid, devic_compris, voltag_batteri, fuel_pump, lower_bumper, agent_control, batteri_electr, power_manag, wheel_drive
FREX: high_voltag, oper_motor, control_motor, power_control, voltag_batteri, lower_bumper, agent_control, batteri_electr, power_manag, wheel_drive, compon_motor, drive_behavior, notif_devic, vibrat_damper, power_generat
Score: high_voltag, steer_wheel, control_motor, power_control, voltag_batteri, oper_motor, light_guid, mobil_devic, wheel_drive, sourc_estim, generat_locat, abnorm_behavior, drive_behavior, agent_control, behavior_notif
Topic 8 Top Words:
Highest Prob: comput_readabl, display_devic, readabl_medium, record_medium, transitori_comput, drive_devic, manag_devic, protect_devic, devic_display, readabl_storag, readabl_record, motor_correspond, devic_comput, medium_control, lock_devic
FREX: comput_readabl, display_devic, readabl_medium, transitori_comput, drive_devic, manag_devic, protect_devic, readabl_record, cup_holder, motor_display, rear_door, devic_manag, manag_program, pressur_vessel, air_condition
Score: comput_readabl, display_devic, readabl_medium, transitori_comput, record_medium, drive_devic, readabl_record, manag_devic, devic_display, readabl_storag, protect_devic, cup_holder, motor_display, ion_batteri, steer_wire
Topic 9 Top Words:
Highest Prob: comput_program, autonom_drive, driver_assist, control_hybrid, shock_absorb, devic_storag, drive_control, hybrid_electr, detect_devic, oper_manag, power_transmiss, pre_chamber, program_product, devic_comput, monitor_devic
FREX: autonom_drive, driver_assist, control_hybrid, drive_control, hybrid_electr, oper_manag, power_transmiss, pre_chamber, hold_devic, ventil_devic, devic_autonom, lane_chang, self_drive, control_engin, fuel_vapor
Score: comput_program, autonom_drive, driver_assist, drive_control, control_hybrid, shock_absorb, hybrid_electr, oper_manag, pre_chamber, devic_storag, road_obstacl, obstacl_detect, fuel_vapor, self_drive, lane_chang
Topic 10 Top Words:
Highest Prob: control_devic, devic_control, learn_devic, estim_devic, data_analysi, devic_power, analysi_devic, abnorm_detect, detect_devic, control_data, damp_control, motor_control, fuel_suppli, heat_exchang, air_fuel
FREX: control_devic, devic_control, learn_devic, control_data, damp_control, air_fuel, control_storag, devic_learn, diagnosi_devic, fuel_ratio, generat_control, control_learn, softwar_updat, devic_hybrid, front_lower
Score: control_devic, devic_control, data_analysi, analysi_devic, control_data, learn_devic, damp_control, fuel_ratio, air_fuel, control_learn, devic_learn, abnorm_detect, generat_control, ratio_sensor, devic_power
Topic 11 Top Words:
Highest Prob: combust_engin, inform_devic, devic_inform, devic_combust, engin_control, object_detect, artifici_muscl, batteri_motor, detect_devic, inform_program, two_stroke, misfir_detect, control_combust, stroke_uniflow, drive_forc
FREX: combust_engin, devic_inform, engin_control, object_detect, artifici_muscl, two_stroke, misfir_detect, control_combust, stroke_uniflow, drive_forc, uniflow_scaveng, engin_oper, center_consol, batteri_case, oper_combust
Score: combust_engin, devic_combust, inform_devic, devic_inform, two_stroke, engin_control, artifici_muscl, misfir_detect, stroke_uniflow, uniflow_scaveng, object_detect, control_combust, oper_combust, larg_turbocharg, inform_program
Topic 12 Top Words:
Highest Prob: high_pressur, motor_motor, pressur_tank, automat_transmiss, tank_manufactur, neural_network, devic_communic, transmiss_control, seat_seat, manufactur_high, storag_compart, least_partial, front_grill, batteri_control, cylind_head
FREX: high_pressur, pressur_tank, automat_transmiss, tank_manufactur, transmiss_control, manufactur_high, storag_compart, front_grill, batteri_control, cylind_head, inject_mold, manufactur_devic, support_element, occup_protect, charg_batteri
Score: pressur_tank, high_pressur, automat_transmiss, motor_motor, manufactur_high, tank_manufactur, storag_compart, transmiss_control, neural_network, inject_mold, cylind_head, front_grill, restraint_devic, tank_high, devic_communic
Topic 13 Top Words:
Highest Prob: storag_medium, transitori_storag, control_control, control_program, inform_inform, medium_store, rear_bumper, electr_motor, traction_batteri, determin_devic, bumper_cover, devic_determin, drive_support, inform_manag, trim_element
FREX: transitori_storag, inform_inform, medium_store, rear_bumper, electr_motor, bumper_cover, drive_support, inform_manag, trim_element, engin_start, balanc_train, store_program, communic_communic, relay_devic, carbon_dioxid
Score: storag_medium, inform_inform, transitori_storag, control_control, medium_store, rear_bumper, control_program, balanc_train, bumper_cover, electr_motor, drive_support, relay_devic, train_control, store_program, traction_batteri

6.2 R&D Uncertainty Matrix

6.2.1 Uncertainty Score Questionnaire

Figure 8

Questionnaire on Market Uncertainty from MacMillan & McGrath (2002)

How certain is YOUR TEAM of the following? Score on scale of 1 (certain) to 7 (highly uncertain).*
Market demand for future products using the fruits of the project
Total future revenues from these products
The stability of the revenue stream generated
Extent to which you will be able to obtain needed support from distributors and suppliers
Extent to which premium pricing can be expected
Extent to which premium pricing can be sustained
The speed with which products will be accepted in the market
The speed with which products will be approved by necessary regulatory bodies
Who the major competitors will be
The probability that competitors will rapidly imitate us
The probability of other technologies matching our offerings
The probability of having our technology blocked by others
Whether the technology has the potential to be licensed
Degree to which we will have to constantly change designs
The degree to which parallel technologies will be needed
Whether such parallel technologies will be available in time
Degree to which technical specifications will be required in the industry
Degree to which technical specifications will be standardized in the industry
The probability of profits being disrupted by third-party intervention (governments, distribution channels, labor unions, etc.)

*Do not answer where you do not know.

Figure 9

Questionnaire on Technological Uncertainty from MacMillan & McGrath (2002)

How certain is YOUR MANAGEMENT TEAM of the following? Score on scale of 1 (certain) to 7 (highly uncertain).*
The time it will take to complete development
The type of skills needed for development
The availability of necessary skills
The cost of staffing those skills
The type of equipment needed for development
The availability of equipment needed
The cost of equipment that is needed
The systems needed for development
The availability of systems needed
The cost of systems needed
The raw materials that will be needed
The availability of needed raw materials
The cost of raw materials
Total costs of development
The infrastructure that needs to be created
Our ability to access necessary complementary technologies
The cost of access to needed complementary technologies
The technology barriers we will face
Our ability to overcome technology barriers we will face
The cost to overcome technology barriers
The required level of product quality
Required levels of support and service
How much production capacity will be needed
The commitment level of senior management

6.2.2 Scoring of Topics

Table 31

Market Uncertainty Scores of Topics

	Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**	**Topic 6**	**Topic 7**	**Topic 8**	**Topic 9**	**Topic 10**	**Topic 11**
Question 1	1	1	2	2	6	NA	3	1	6	3	3
Question 2	3	1	2	3	7	NA	3	NA	6	2	2
Question 3	2	1	3	2	6	NA	4	NA	4	1	1
Question 4	1	1	3	2	7	3	3	2	6	2	1
Question 5	2	3	2	4	7	2	4	3	6	2	1
Question 6	2	4	5	4	7	2	6	3	6	2	1
Question 7	3	1	5	3	5	1	NA	1	6	1	1
Question 8	1	1	3	3	7	1	NA	1	7	1	1
Question 9	1	2	3	2	6	3	1	2	5	1	1
Question 10	1	2	1	2	5	4	3	1	5	2	1
Question 11	3	3	3	3	5	7	5	1	6	1	2
Question 12	1	2	3	2	6	1	1	1	7	2	1
Question 13	3	2	1	1	5	4	3	2	5	1	1
Question 14	1	2	2	2	5	1	4	2	5	2	1
Question 15	2	2	3	3	5	1	3	1	6	1	2
Question 16	2	5	5	5	6	2	3	2	6	2	2
Question 17	2	2	1	2	5	1	2	1	5	1	1
Question 18	5	3	3	4	6	1	2	1	5	1	1
Question 19	3	3	3	3	4	2	2	3	6	1	3
Average Score	**2,1**	**2,2**	**2,8**	**2,7**	**5,8**	**2,3**	**3,1**	**1,6**	**5,7**	**1,5**	**1,4**

Table 32

Technological Uncertainty Scores of Topics

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11
Question 1	6	5	6	4	5	2	5	3	6	2	3
Question 2	3	2	3	6	5	1	5	1	5	1	2
Question 3	1	2	3	6	3	1	5	1	5	1	1
Question 4	3	3	4	7	4	2	4	2	6	1	1
Question 5	2	2	3	5	4	1	3	1	7	1	2
Question 6	1	1	3	5	3	1	2	1	6	1	1
Question 7	3	2	5	7	5	1	3	3	6	1	2
Question 8	3	3	4	6	6	1	5	2	5	1	2
Question 9	1	2	3	6	4	1	5	1	5	1	1
Question 10	3	3	3	5	6	2	5	3	6	2	2
Question 11	2	1	3	4	4	3	6	1	3	1	1
Question 12	2	1	4	4	4	3	6	2	3	2	1
Question 13	2	1	4	6	5	4	4	2	3	2	1
Question 14	5	7	6	7	7	4	6	4	7	2	4
Question 15	6	7	4	6	3	2	5	3	7	3	3
Question 16	2	4	3	6	3	1	3	2	6	2	1
Question 17	3	4	5	5	4	3	5	3	6	2	2
Question 18	5	5	4	6	5	2	4	2	7	2	2
Question 19	3	3	5	5	5	2	4	2	6	1	3
Question 20	6	7	7	7	7	4	6	5	7	3	3
Question 21	2	2	1	5	2	3	1	3	2	2	1
Question 22	1	3	6	5	5	3	3	3	5	2	3
Question 23	3	2	6	5	6	1	2	3	5	1	1
Question 24	3	1	1	3	3	4	2	3	2	3	2
Average Score	**3,0**	**3,0**	**4,0**	**5,5**	**4,5**	**2,2**	**4,1**	**2,3**	**5,3**	**1,7**	**1,9**

6.3 R Syntax

```
#installing and loading packages
install.packages("tm")
install.packages("readxl")
install.packages("quanteda")
install.packages("ldatuning")
install.packages('stm')
install.packages('dplyr')
library("stm")
library("tm")
library("readxl")
library("quanteda")
library("ldatuning")
library("dplyr")
#setting file directory and reading in excel dataset
setwd("C:/Users/User/Documents/RStudio")
path_data<- system.file("extdata/", package = "readtext")
data<- read_xlsx(paste0(path_data, "lookup.xlsx"))
#creating the corpus and performing pre-processing
corpus<-corpus(data, text_field = "PatentTitle")
tokenized<-tokens(corpus, remove_punct = TRUE, remove_numbers = TRUE, remove_symbols = TRUE,
remove_separators = TRUE, split_hyphens = TRUE)
tokenized<-tokens_keep(tokenized, pattern = "^[a-zA-Z]+$", valuetype = "regex")
tokenized<-tokens_select(tokenized, pattern = stopwords("en"), selection = "remove")
tokenized<-tokens_tolower(tokenized)
```

```
tokenized<-tokens_wordstem(tokenized)
tokenized<-tokens_remove(tokenized, c('method','vehicl','car','system', 'process', 'assembl', 'apparatus', 'non', 'use', 'thereof','vehicular','intern','unit', 'toy', 'replica'))
tokenized<-tokens_ngrams(tokenized, n = 2)
#creating dfm, subsetting it and trimming it, lastly dropping empty rows
dfm<-dfm(tokenized)
dfm<-dfm_subset(dfm, ntoken(dfm)>0)
dfm2020<-dfm_subset(dfm, Year >= 2020)
dfm2020<-dfm2020 %>%
  dfm_trim(min_termfreq = 0.8, termfreq_type = "quantile", max_docfreq = 0.1, docfreq_type = "prop")
dfm2020<-dfm_subset(dfm2020, ntoken(dfm2020)>0)
#finding optimal topic number and plotting it
ldatest<-FindTopicsNumber(dfm2020, topics= seq(from = 2, to = 200, by = 2 ), metrics = c( "CaoJuan2009", "Arun2010","Griffiths2004", "Deveaud2014"), method = "Gibbs", control = list(seed = 77), verbose = TRUE)
FindTopicsNumber_plot(ldatest)
#fitting the lda from the dfm
stmmodel2020<-stm(dfm2020, K = 13, seed = 123, init.type = 'LDA')
#commanding label output
labelTopics(stmmodel2020, n= 10)
#plotting frequencies
plot(stmmodel2020)
```

6.4 List of R Packages

The packages used were: stm, tm, readxl, quanteda, ldatuning, dplyr

7. References

Aaldering, L. J., Leker, J., & Song, C. H. (2019). Competition or collaboration? – Analysis of technological knowledge ecosystem within the field of alternative powertrain systems: A patent-based approach. *Journal of Cleaner Production*, *212*, 362–371. https://doi.org/10.1016/J.JCLEPRO.2018.12.047

Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of Artificial Intelligence in Transport: An Overview. *Sustainability 2019, Vol. 11, Page 189*, *11*(1), 189. https://doi.org/10.3390/SU11010189

Aichholzer, G. (2009). The Delphi Method: Eliciting Experts' Knowledge in Technology Foresight. *Interviewing Experts*, 252–274. https://doi.org/10.1057/9780230244276_13

Allen, S. D., Sarkis -, J., Agboola, M. G., Awobajo, K. A., Oluwatobi, S. O., -, al, & Ignat, V. (2017). Digitalization and the global technology trends. *IOP Conference Series: Materials Science and Engineering*, *227*(1), 012062. https://doi.org/10.1088/1757-899X/227/1/012062

Antons, D., Grünwald, E., Cichy, P., & Salge, T. O. (2020). The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R&D Management*, *50*(3), 329–351. https://doi.org/10.1111/RADM.12408

Apreda, R., Bonaccorsi, A., Dell'orletta, F., & Fantoni, G. (2016). Functional technology foresight. A novel methodology to identify emerging technologies. *European Journal of Futures Research*, *4*(1), 1–15. https://doi.org/10.1007/S40309-016-0093-1/TABLES/4

Archibugi, D. (1992). Patenting as an indicator of technological innovation: a review. *Science and Public Policy*, *19*(6), 357–368. https://doi.org/10.1093/SPP/19.6.357

Arun, R., Suresh, V., Madhavan, C. E. V., & Murty, M. N. (2010). On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6118 LNAI*(PART 1), 391–402. https://doi.org/10.1007/978-3-642-13657-3_43

Asif, U., & Schmidt, K. (2021). Fuel Cell Electric Vehicles (FCEV): Policy Advances to Enhance Commercial Success. *Sustainability 2021, Vol. 13, Page 5149*, *13*(9), 5149. https://doi.org/10.3390/SU13095149

Askaripoor, H. ;, Hashemi, M. ;, Knoll, A. E., Ruiz, J. A., Llamazares, A., Lauer, M., Askaripoor, H., Farzaneh, M. H., & Knoll, A. (2022). *E/E Architecture Synthesis: Challenges and Technologies*. https://doi.org/10.3390/electronics11040518

Audi AG. (2021). *Ein Claim mit Geschichte: Audi feiert 50 Jahre „Vorsprung durch Technik"*. Audi Mediacenter.

Bennett, N., & Lemoine, J. (2014). What VUCA Really Means for You. *Harvard Business Review*, *92*(2), 9–22. https://papers.ssrn.com/abstract=2389563

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, *3*(30), 774. https://doi.org/10.21105/joss.00774

Bischof, J. M., & Airoldi, E. M. (2012). Summarizing Topical Content with Word Frequency and Exclusivity. *Proceedings of the 29th International Coference on International Conference on Machine Learning*, 9–16.

Bouwman, H., Akademi, Å., Carlsson, F. C., Carlsson, J., Nikou, S., Sell, A., & Walden, P. (2014). *How Nokia failed to nail the Smartphone market*. https://www.econstor.eu/handle/10419/101414

Burcicki, D. (2020). *The E/E architecture and the future of automotive innovation - Capital*. Siemens Blog. https://blogs.sw.siemens.com/ee-systems/2020/06/09/the-ee-architecture-and-the-future-of-automotive-innovation/

Bureau van Dijk. (2023). *Orbis*. Bureau van Dijk.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience 2013 14:5*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, *72*(7–9), 1775–1781. https://doi.org/10.1016/J.NEUCOM.2008.06.011

Chang, J. (2015). *lda: Collapsed Gibbs Sampling Methods for Topic Models.* https://cran.r-project.org/web/packages/lda/index.html

Cifci, H., & Yuksel, N. (2018). Foresight 6.0: The New Generation of Technology Foresight. *2018 IEEE International Conference on Engineering, Technology and Innovation, ICE/ITMC 2018 - Proceedings*. https://doi.org/10.1109/ICE.2018.8436350

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (1999). New product portfolio management: practices and performance. *Journal of Product Innovation Management*, *16*(4), 333–351. https://doi.org/10.1016/S0737-6782(99)00005-3

Danieli, P., Masi, M., Lazzaretto, A., Carraro, G., Dal Cin, E., & Volpato, G. (2023). Is Banning Fossil-Fueled Internal Combustion Engines the First Step in a Realistic Transition to a 100% RES Share? *Energies 2023, Vol. 16, Page 5690*, *16*(15), 5690. https://doi.org/10.3390/EN16155690

de Miranda Santo, M., 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. https://doi.org/10.1016/J.TECHFORE.2006.05.020

Deveaud, R., Sanjuan, E., & Bellot, P. (2014). *Accurate and Effective Latent Concept Modeling for Ad Hoc Information Retrieval*. https://doi.org/10.3166/dn.17.1.61-84

Dioha, M. O., Lukuyu, J., Virgüez, E., & Caldeira, K. (2022). Guiding the deployment of electric vehicles in the developing world. *Environ. Res. Lett*, *17*, 71001. https://doi.org/10.1088/1748-9326/ac765b

Dixon, R. (2023, April 22). *Be ready for the coming shift in automotive computing power | S&P Global*. S&P Global Mobility Blog. https://www.spglobal.com/mobility/en/research-analysis/be-ready-for-the-coming-shift-in-automotive-computing-power.html

Drew, S. A. W. (2006). Building technology foresight: Using scenarios to embrace innovation. *European Journal of Innovation Management*, *9*(3), 241–257. https://doi.org/10.1108/14601060610678121/FULL/PDF

Eckert, T., & Hüsig, S. (2021). Innovation portfolio management: a systematic review and research agenda in regards to digital service innovations. *Management Review Quarterly 2021 72:1*, *72*(1), 187–230. https://doi.org/10.1007/S11301-020-00208-3

European Commission. (2022). *Research and development expenses and intensity of selected global automotive manufacturers in 2021*. https://www-statista-com.ezproxy2.utwente.nl/statistics/1306984/randd-in-selected-global-automotive-oems-expenses-and-intensity/

European Environment Agency. (2016). *Fleet composition*. https://www.eea.europa.eu/publications/ENVISSUENo12/page032.html

Gerhard, D., Brem, A., & Voigt, K. I. (2008). Product development in the automotive industry: Crucial success drivers for technological innovations. *International Journal of Technology Marketing*, *3*(3), 203–222. https://doi.org/10.1504/IJTMKT.2008.019922

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(SUPPL. 1), 5228–5235. https://doi.org/10.1073/PNAS.0307752101/ASSET/A171028F-351B-4832-8A36-1B0244146E12/ASSETS/GRAPHIC/ZPQ0080438990006.JPEG

Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, *17*, 26–32. https://doi.org/10.1016/J.PROCS.2013.05.005

Hu, R., Ma, W., Lin, W., Chen, X., Zhong, Z., & Zeng, C. (2022). *Technology Topic Identification and Trend Prediction of New Energy Vehicle Using LDA Modeling*. https://doi.org/10.1155/2022/9373911

Hussain, M., Tapinos, E., & Knight, L. (2017). Scenario-driven roadmapping for technology foresight. *Technological Forecasting and Social Change*, *124*, 160–177. https://doi.org/10.1016/J.TECHFORE.2017.05.005

Iamratanakul, S., Shankar, R., & Dimmitt, N. J. (2009). Improving project portfolio management with strategic alignment. *PICMET: Portland International Center for Management of Engineering and Technology, Proceedings*, 1290–1300. https://doi.org/10.1109/PICMET.2009.5261993

IBIS World. (2022). *Global automotive manufacturing market size 2022 | Statista*. https://www-statista-com.ezproxy2.utwente.nl/statistics/574151/global-automotive-industry-revenue/

IbisWorld. (2023). *Global car manufacturing industry revenue between 2019 and 2022, with a forecast for 2023* . https://www-statista-com.ezproxy2.utwente.nl/statistics/574151/global-automotive-industry-revenue/

Ilevbare, I. M., Probert, D., & Phaal, R. (2013). A review of TRIZ, and its benefits and challenges in practice. *Technovation*, *33*(2–3), 30–37. https://doi.org/10.1016/J.TECHNOVATION.2012.11.003

Isermann, R. (2001). Diagnosis methods for electronic controlled vehicles. *Vehicle System Dynamics*, *36*(2–3), 77–117. https://doi.org/10.1076/VESD.36.2.77.3554

Jalonen, H. (2011). The Uncertainty of Innovation: A Systematic Review of the Literature Co-creation of Service Innovations in Europe (CoSIE) View project The Uncertainty of Innovation: A Systematic Review of the Literature. *Article in Journal of Management Research*, *4*(1), 12. https://doi.org/10.5296/jmr.v4i1.1039

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/S11042-018-6894-4/TABLES/11

Jugend, D., & Da Silva, S. L. (2014). Product-portfolio management: A framework based on methods, organization, and strategy. *Concurrent Engineering Research and Applications*, *22*(1), 17–28. https://doi.org/10.1177/1063293X13508660/ASSET/IMAGES/LARGE/10.1177_1063293X13508660-FIG2.JPEG

Junquera, B., Moreno, B., & Álvarez, R. (2016). Analyzing consumer attitudes towards electric vehicle purchasing intentions in Spain: Technological limitations and vehicle confidence. *Technological Forecasting and Social Change*, *109*, 6–14. https://doi.org/10.1016/J.TECHFORE.2016.05.006

Kanama, D., Kondo, A., & Yokoo, Y. (2008). Development of technology foresight: Integration of technology roadmapping and the Delphi method. *International Journal of Technology Intelligence and Planning*, *4*(2), 184–200. https://doi.org/10.1504/IJTIP.2008.018316

Kayser, V., & Blind, K. (2017a). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*, *116*, 208–215. https://doi.org/10.1016/J.TECHFORE.2016.10.017

Kayser, V., & Blind, K. (2017b). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*, *116*, 208–215. https://doi.org/10.1016/J.TECHFORE.2016.10.017

Kayser, V., Goluchowicz, K., & Bierwisch, A. (2014). TEXT MINING FOR TECHNOLOGY ROADMAPPING — THE STRATEGIC VALUE OF INFORMATION. *Https://Doi.Org/10.1142/S1363919614400040*, *18*(3). https://doi.org/10.1142/S1363919614400040

Kehoe, C. A., & Yu, X. J. (2001). Patent Data for Technology Assessment, Part I. *Https://Doi.Org/10.1300/J122v22n01_07*, 101–116. https://doi.org/10.1300/J122V22N01_07

Kim, K. H., Han, Y. J., Lee, S., Cho, S. W., & Lee, C. (2019). Text Mining for Patent Analysis to Forecast Emerging Technologies in Wireless Power Transfer. *Sustainability 2019, Vol. 11, Page 6240*, *11*(22), 6240. https://doi.org/10.3390/SU11226240

Klein, C., Hoj, J., & Machlica, G. (2021). *The impacts of the COVID-19 crisis on the automotive sector in Central and Eastern European Countries | OECD Economics Department Working Papers | OECD iLibrary*. https://www.oecd-ilibrary.org/economics/the-impacts-of-the-covid-19-crisis-on-the-automotive-sector-in-central-and-eastern-european-countries_a7d40030-en

Kolominsky-Rabas, P. L., Djanatliev, A., Wahlster, P., Gantner-Bär, M., Hofmann, B., German, R., Sedlmayr, M., Reinhardt, E., Schüttler, J., Kriza, C., Niederländer, C., Prokosch, H. U., Lenz, R., Baumgärtel, P., Schöffski, O., Emmert, M., Meier, F., Aisenbrey, A., Voigt, W., … Miethe, M. (2015). Technology foresight for medical device development through hybrid simulation: The ProHTA Project. *Technological Forecasting and Social Change*, *97*, 105–114. https://doi.org/10.1016/J.TECHFORE.2013.12.005

Koppel, S., Charlton, J., Fildes, B., & Fitzharris, M. (2008). How important is vehicle safety in the new vehicle purchase process? *Accident Analysis & Prevention*, *40*(3), 994–1004. https://doi.org/10.1016/J.AAP.2007.11.006

Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S., Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E., Palacios-Marqués, D., Schiavone, F., Ferraris, A., Fernandes, C., & Ferreira, J. J.

(2022). Literature reviews as independent studies: guidelines for academic practice. *Review of Managerial Science*, *16*(8), 2577–2595. https://doi.org/10.1007/S11846-022-00588-8/FIGURES/3

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 419–431. https://doi.org/10.1016/J.TECHFORE.2012.10.005

Lang, A., & Murphy, H. (2014). *Business and Sustainability: An Introduction*. 3–19. https://doi.org/10.1007/978-3-319-07239-5_1

Larreina, S., & Hernando, S. (2005). Application Of Technology Prospective To Business Sectorial Studies. *WIT Transactions on Information and Communication Technologies*, *35*, 321–329. https://doi.org/10.2495/DATA050341

Lee, H., & Kang, P. (2018). Identifying core topics in technology and innovation management studies: a topic model approach. *Journal of Technology Transfer*, *43*(5), 1291–1317. https://doi.org/10.1007/S10961-017-9561-4/METRICS

Lee, J. Y., Kumari, R., Jeong, J. Y., Kim, T. H., & Lee, B. H. (2020). Knowledge Discovering on Graphene Green Technology by Text Mining in National R&D Projects in South Korea. *Sustainability 2020, Vol. 12, Page 9857*, *12*(23), 9857. https://doi.org/10.3390/SU12239857

Liu, Y., Li, K. J., Chen, H. A., & Balachander, S. (2017). The Effects of Products' Aesthetic Design on Demand and Marketing-Mix Effectiveness: The Role of Segment Prototypicality and Brand Consistency. *Https://Doi.Org/10.1509/Jm.15.0315*, *81*(1), 83–102. https://doi.org/10.1509/JM.15.0315

Llopis-Albert, C., Rubio, F., & Valero, F. (2021). Impact of digital transformation on the automotive industry. *Technological Forecasting and Social Change*, *162*. https://doi.org/10.1016/J.TECHFORE.2020.120343

Lukin, E., Krajnović, A., & Bosna, J. (2022). Sustainability Strategies and Achieving SDGs: A Comparative Analysis of Leading Companies in the Automotive Industry. *Sustainability 2022, Vol. 14, Page 4000*, *14*(7), 4000. https://doi.org/10.3390/SU14074000

Luo, L. M. (2011). Optimal diversification for R&D project portfolios. *Scientometrics*, *91*(1), 219–229. https://doi.org/10.1007/S11192-011-0537-0

Mack, O., & Khare, A. (2015). Perspectives on a VUCA world. *Managing in a VUCA World*, 3–19. https://doi.org/10.1007/978-3-319-16889-0_1/COVER

MacMillan, I. C., & McGrath, R. G. (2002). Crafting R&D Project Portfolios. *Http://Dx.Doi.Org/10.1080/08956308.2002.11671522*, *45*(5), 48–59. https://doi.org/10.1080/08956308.2002.11671522

Madani, F., & Weber, C. (2016). The evolution of patent mining: Applying bibliometrics analysis and keyword network analysis. *World Patent Information*, *46*, 32–48. https://doi.org/10.1016/J.WPI.2016.05.008

Madhok, A., & Osegowitsch, T. (2000). The international biotechnology industry: A dynamic capabilities perspective. *Journal of International Business Studies*, *31*(2), 325–335. https://doi.org/10.1057/PALGRAVE.JIBS.8490909/METRICS

Magruk, A. (2011). Innovative classification of technology foresight methods. *Technological and Economic Development of Economy*, *17*(4), 700–715. https://doi.org/10.3846/20294913.2011.649912

Mallick, D. N., & Schroeder, R. G. (2005). An integrated framework for measuring product development performance in high technology industries. *PRODUCTION AND OPERATIONS MANAGEMENT*, *14*(2), 142–158.

Martini, G., Astorga, C., Adam, T., Bonnel, P., Farlafetti, A., Junninen, H., Manfredi, U., Montero, L., Müller, A., Krasenbrink, A., Larsen, B., Rey, M., & De Santi, G. (2009). *Physical & Chemical Characterization of emissions from 2-Stroke motorcycles*. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC53779/eur%20report%2023999%20en_mopeds_2009%20%282%29.pdf

Marutho, D., Hendra Handaka, S., Wijaya, E., & Muljono. (2018). The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISemantic 2018*, 533–538. https://doi.org/10.1109/ISEMANTIC.2018.8549751

Meidute-Kavaliauskiene, I., Yıldız, B., Semsettin, ¸, Gdem, Ç. ˘, Renatačinčikait˙, R. R., Yair, W., & Raudeliuniene, J. (2021). Do People Prefer Cars That People Don't Drive? A Survey Study on Autonomous Vehicles. *Energies 2021, Vol. 14, Page 4795*, *14*(16), 4795. https://doi.org/10.3390/EN14164795

Mercedes-Benz Group AG. (2022). *Mercedes-Benz Group Annual Report 2022*. https://group.mercedes-benz.com/documents/investors/reports/annual-report/mercedes-benz/mercedes-benz-annual-report-2022-incl-combined-management-report-mbg-ag.pdf

Miles, I. (2010). The development of technology foresight: A review. *Technological Forecasting and Social Change*, *77*(9), 1448–1456. https://doi.org/10.1016/J.TECHFORE.2010.07.016

Minghui, Z., Hanrui, Y., Yao, P., & Lingling, Z. (2022). Literature Review and Practice Comparison of Technology Foresight. *Procedia Computer Science*, *199*, 837–844. https://doi.org/10.1016/J.PROCS.2022.01.104

Mitchell, G. R., & Hamilton, W. F. (1988). Managing R&D As a Strategic Option. *Research-Technology Management*, *31*(3), 15–22. https://doi.org/10.1080/08956308.1988.11670521

Moro, S., Cortez, P., & Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Systems with Applications*, *42*(3), 1314–1324. https://doi.org/10.1016/J.ESWA.2014.09.024

Mugge, R., & Dahl, D. W. (2013). Seeking the Ideal Level of Design Newness: Consumer Response to Radical and Incremental Product Design. *Journal of Product Innovation Management*, *30*(SUPPL 1), 34–47. https://doi.org/10.1111/JPIM.12062

Mühlroth, C., & Grottke, M. (2018). A systematic literature review of mining weak signals and trends for corporate foresight. *Journal of Business Economics 2018 88:5*, *88*(5), 643–687. https://doi.org/10.1007/S11573-018-0898-4

Murry, J. W., & Hammons, J. O. (1995). Delphi: A Versatile Methodology for Conducting Qualitative Research. *The Review of Higher Education*, *18*(4), 423–436. https://doi.org/10.1353/RHE.1995.0008

Murzintcev, N. (2020). *ldatuning: Tuning of the Latent Dirichlet Allocation Models Parameters*. https://CRAN.R-project.org/package=ldatuning

Nayak, J., Mishra, | Manohar, Naik, | Bighnaraj, Swapnarekha, H., Korhan Cengiz, |, & Shanmuganathan, V. (2021). *An impact study of COVID-19 on six different industries: Automobile, energy and power, agriculture, education, travel and tourism and consumer electronics*. https://doi.org/10.1111/exsy.12677

Neudorfer, A., & Voos, W. (1977). PATENT ANALYSIS, CONDUCTED SYSTEMATICALLY WITH REFERENCE TO SKI BINDINGS. *F&M-FEINWERKTECHNIK & MESSTECHNIK*, *85*(8), 385–392.

OECD Publishing. (2012). *Main Science and Technology Indicators*.

Office for Official Publications of the European Communities. (2008). *NACE Rev. 2 - Statistical classification of economic activities*. https://www.destatis.de/Europa/EN/Publications/General-regional-statistics/MW_KSRA07015EN.html

Ojapah, M. M., Zhang, Y., & Zhao, H. (2013). Part-load performance and emissions analysis of SI combustion with EIVC and throttled operation and CAI combustion. *Internal Combustion Engines: Performance, Fuel Economy and Emissions: IMechE, London, 27-28 November 2013*, 19–32. https://doi.org/10.1533/9781782421849.1.19

Oliveira, L. P. De, Wehrmeister, M. A., & Oliveira, A. S. De. (2017). Systematic Literature Review on Automotive Diagnostics. *Brazilian Symposium on Computing System Engineering, SBESC*, *2017-November*, 1–8. https://doi.org/10.1109/SBESC.2017.7

Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: purpose, process, and structure. *Journal of the Academy of Marketing Science*, *46*(1), 1–5. https://doi.org/10.1007/S11747-017-0563-4/METRICS

Paquin, J. P., Tessier, D., & Gauthier, C. (2015). The Effectiveness of Portfolio Risk Diversification: An Additive Approach by Project Replication. *Https://Doi.Org/10.1002/Pmj.21526*, *46*(5), 94–110. https://doi.org/10.1002/PMJ.21526

Park, J. S., Kim, N. R., Choi, H.-R., Han, E., Park, J. S., Kim, N. R., Choi, H.-R., & Han, E. (2018). A New Forecasting System Using the Latent Dirichlet Allocation (LDA) Topic Modeling Technique. *WSEAS Transactions on Environment and Development*, *14*, 363–373. https://www.wseas.com/journals/articles.php?id=2585

Pelliccione, P., Knauss, E., Heldal, R., Magnus Ågren, S., Mallozzi, P., Alminger, A., & Borgentun, D. (2017). Automotive Architecture Framework: The experience of Volvo Cars. *Journal of Systems Architecture*, *77*, 83–100. https://doi.org/10.1016/J.SYSARC.2017.02.005

Pietrobelli, C., & Puppato, F. (2016). Technology foresight and industrial strategy. *Technological Forecasting and Social Change*, *110*, 117–125. https://doi.org/10.1016/J.TECHFORE.2015.10.021

Pope, C., Ziebland, S., & Mays, N. (2000). Analysing qualitative data. *BMJ*, *320*(7227), 114–116. https://doi.org/10.1136/BMJ.320.7227.114

Porter, A. L. (2010). Technology foresight: Types and methods. *International Journal of Foresight and Innovation Policy*, *6*(1–3), 36–45. https://doi.org/10.1504/IJFIP.2010.032664

Porter, A. L. (2015). How “Tech Mining” Can Enhance R&D Management. *Http://Dx.Doi.Org/10.1080/08956308.2007.11657425*, *50*(2), 15–20. https://doi.org/10.1080/08956308.2007.11657425

Qian, Z. Q., Siegmann, K., Keller, A., Matter, U., Scherrer, L., & Siegmann, H. C. (2000). Nanoparticle air pollution in major cities and its origin. *ATMOSPHERIC ENVIRONMENT*, *34*(3), 443–451.

R&D World. (2022). *Global automotive R&D spending | Statista*. https://www-statista-com.ezproxy2.utwente.nl/statistics/1345699/global-automotive-research-and-development-spending/

Reger, G. (2001). Technology foresight in companies: From an indicator to a network and process perspective. *Technology Analysis and Strategic Management*, *13*(4), 533–553. https://doi.org/10.1080/09537320120095437

Reuter, C., Foerstl, K., Hartmann, E., & Blome, C. (2010). SUSTAINABLE GLOBAL SUPPLIER MANAGEMENT: THE ROLE OF DYNAMIC CAPABILITIES IN ACHIEVING COMPETITIVE ADVANTAGE. *Journal of Supply Chain Management*, *46*(2), 45–63. https://doi.org/10.1111/J.1745-493X.2010.03189.X

Riesener, M., Kuhn, M., Lauf, H., Manoharan, S., & Schuh, G. (2022). Concept for the identification of product innovation potentials by the application of text mining. *Procedia CIRP*, *109*, 281–286. https://doi.org/10.1016/J.PROCIR.2022.05.250

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R Package for Structural Topic Models. *Journal of Statistical Software*, *91*(2), 1–40. https://doi.org/10.18637/jss.v091.i02

Rogers, Everett M, Singhal, A., & Quinlan, M. M. (2014). Diffusion of Innovations. *Taylor & Francis Group*, 432–448. https://doi.org/10.4324/9780203887011-36

Saibannavar, D., Math, M. M., & Kulkarni, U. (2021). A Survey on On-Board Diagnostic in Vehicles. *EAI/Springer Innovations in Communication and Computing*, 49–60. https://doi.org/10.1007/978-3-030-49795-8_5/COVER

Shih, M. J., Liu, D. R., & Hsu, M. L. (2010). Discovering competitive intelligence by mining changes in patent trends. *Expert Systems with Applications*, *37*(4), 2882–2890. https://doi.org/10.1016/J.ESWA.2009.09.001

Sinigaglia, T., Eduardo Santos Martins, M., & Cezar Mairesse Siluk, J. (2022). Technological evolution of internal combustion engine vehicle: A patent data analysis. *Applied Energy*, *306*, 118003. https://doi.org/10.1016/J.APENERGY.2021.118003

Sirmon, D. G., Hitt, M. A., & Ireland, R. D. (2007). Managing Firm Resources in Dynamic Environments to Create Value: Looking Inside the Black Box. *Https://Doi.Org/10.5465/Amr.2007.23466005*, *32*(1), 273–292. https://doi.org/10.5465/AMR.2007.23466005

Song, C. H., & Aaldering, L. J. (2019). Strategic intentions to the diffusion of electric mobility paradigm: The case of internal combustion engine vehicle. *Journal of Cleaner Production*, *230*, 898–909. https://doi.org/10.1016/J.JCLEPRO.2019.05.126

Stellantis N.V. (2023). *Stellantis N.V Annual Report 2022*. https://www.afm.nl/en/sector/registers.

Stoehr, N., Braesemann, F., Frommelt, M., & Zhou, S. (2020). Mining the Automotive Industry: A Network Analysis of Corporate Positioning and Technological Trends. *Springer Proceedings in Complexity*, 297–308. https://doi.org/10.1007/978-3-030-40943-2_25/COVER

Takahashi, S., Owan, H., Tsuru, T., & Uehara, K. (2014). *Discussion Paper Series A No.614 Multitasking Incentives and Biases in Subjective Performance Evaluation*.

Taub, A. I., Krajewski, P. E., Luo, A. A., & Owens, J. N. (2007). Yesterday, today and tomorrow: The evolution of technology for materials processing over the last 50 years: The automotive example. *JOM*, *59*(2), 48–57. https://doi.org/10.1007/S11837-007-0022-7

Tegart, G. (2014). Technology foresight: Philosophy and principles. *Https://Doi.Org/10.5172/Impp.2003.5.2-3.279*, *5*(2–3), 279–285. https://doi.org/10.5172/IMPP.2003.5.2-3.279

Tian, C., Zhang, J., Liu, D., Wang, Q., & Lin, S. (2022). Technological topic analysis of standard-essential patents based on the improved Latent Dirichlet Allocation (LDA) model. *Technology Analysis & Strategic Management*. https://doi.org/10.1080/09537325.2022.2130039

Tong, Z., & Zhang, H. (2016). *A Text Mining Research Based on LDA Topic Modelling*. 201–210. https://doi.org/10.5121/CSIT.2016.60616

Toyota Motor Corporation. (2022). *Annual Report 2022*. https://www.sec.gov/Archives/edgar/data/1094517/000119312522179197/d696693d20f.htm

Van De Kaa, G., Scholten, D., Rezaei, J., & Milchram, C. (2017). The Battle between Battery and Fuel Cell Powered Electric Vehicles: A BWM Approach. *Energies 2017, Vol. 10, Page 1707*, *10*(11), 1707. https://doi.org/10.3390/EN10111707

Vandepaer, L., Cloutier, J., & Amor, B. (2017). Environmental impacts of Lithium Metal Polymer and Lithium-ion stationary batteries. *Renewable and Sustainable Energy Reviews*, *78*, 46–60. https://doi.org/10.1016/J.RSER.2017.04.057

Vílchez-Román, C., Huamán-Delgado, F., & Sanguinetti-Cordero, S. (2019). Topic modeling applied to business research: A latent dirichlet allocation (LDA)-based classification for organization studies. *Communications in Computer and Information Science*, *898*, 212–219. https://doi.org/10.1007/978-3-030-11680-4_21/COVER

Volkswagen AG. (2023). *Key R&D figures - Volkswagen Group Annual Report 2022*. https://annualreport2022.volkswagenag.com/group-management-report/sustainable-value-enhancement/research-and-development/key-r-d-figures.html

Walker, G. H., Stanton, N. A., & Young, M. S. (2001). Where Is Computing Driving Cars? *International Journal of Human-Computer Interaction*, *13*(2), 203–229. https://doi.org/10.1207/S15327590IJHC1302_7

Wang, C., Tu, Z., Chang, L., Zheng, Y., & Zhang, B. (2023). Perspectives on Vehicle Powertrain Control Parameter Calibration Technology. *2023 9th International Conference on Control Science and Systems Engineering, ICCSSE 2023*, 153–158. https://doi.org/10.1109/ICCSSE59359.2023.10245406

Wang, M. Y., Chang, D. S., & Kao, C. H. (2010a). Identifying technology trends for R&D planning using TRIZ and text mining. *R&D Management*, *40*(5), 491–509. https://doi.org/10.1111/J.1467-9310.2010.00612.X

Wang, M. Y., Chang, D. S., & Kao, C. H. (2010b). Identifying technology trends for R&D planning using TRIZ and text mining. *R&D Management*, *40*(5), 491–509. https://doi.org/10.1111/J.1467-9310.2010.00612.X

Wang, W., Feng, Y., & Dai, W. (2018). Topic analysis of online reviews for two competitive products using latent Dirichlet allocation. *Electronic Commerce Research and Applications*, *29*, 142–156. https://doi.org/10.1016/J.ELERAP.2018.04.003

Wellbrock, W., Ludin, D., Röhrle, L., & Gerstlberger, W. (2020). Sustainability in the automotive industry, importance of and impact on automobile interior – insights from an empirical survey. *International Journal of Corporate Social Responsibility 2020 5:1*, *5*(1), 1–11. https://doi.org/10.1186/S40991-020-00057-Z

Woodside, A. G. (2016). The good practices manifesto: Overcoming bad practices pervasive in current research in business. *Journal of Business Research*, *69*(2), 365–381. https://doi.org/10.1016/J.JBUSRES.2015.09.008

World Intellectual Property Organization. (n.d.). *Frequently Asked Questions: Patents*. Retrieved December 18, 2023, from https://www.wipo.int/patents/en/faq_patents.html

Wynn-Williams, M. S. (2009). Economies of scale in the automobile industry. *Surfing the Global Tide*, 41–71. https://doi.org/10.1057/9780230233591_3

Yalcinkaya, G., Calantone, R. J., & Griffith, D. A. (2007). An examination of exploration and exploitation capabilities: Implications for product innovation and market performance. *Journal of International Marketing*, *15*(4), 63–93. https://doi.org/10.1509/JIMK.15.4.63/ASSET/IMAGES/LARGE/10.1509_JIMK.15.4.63-FIG2.JPEG

Zapata, C., & Nieuwenhuis, P. (2010). Exploring innovation in the automotive industry: new technologies for cleaner cars. *Journal of Cleaner Production*, *18*(1), 14–20. https://doi.org/10.1016/J.JCLEPRO.2009.09.009

Zeng, X., Li, M., Abd El-Hady, D., Alshitari, W., Al-Bogami, A. S., Lu, J., & Amine, K. (2019). Commercialization of Lithium Battery Technologies for Electric Vehicles. *Advanced Energy Materials*, *9*(27), 1900161. https://doi.org/10.1002/AENM.201900161