Assessing technology composition of firms in the automotive industry

Author: Sam Gels
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,
In the VUCA world R&D management is continuously confronted with an increase in complexity, speed, and uncertainty in their environment. To battle this, high-tech firms should aim to improve their competitiveness by consistently improving their ability to respond to these changes in their environment and develop compatible strategies to ensure and enhance the organization's survivability. Moreover, Research and Development (R&D) should have the right strategic allocation of all available resources since it takes a long time between the innovation and the commercialization of the R&D project, where pursuing the wrong projects lead to missing out of other opportunities. Therefore, the firms that are better in detecting new technologies trends early-on are more likely to allocate their resources more effectively. This leads us to the aim of the research in this paper, which is to find a reproducible approach for detecting technology trends early-on. This is done in this research with the case study of the automotive industry, where the technology composition of the top 10 most innovative companies based upon their granted patent titles are analysed. In order to protect firms better from the VUCA world this study uses patent analysis with the aim to create a better understanding of emerging technologies and R&D management implications. This study builds upon existing knowledge in the fields of technology trends and R&D portfolio management by introducing an innovative approach that combines text mining and patent analysis. By employing a case study within the automotive industry, this research aims to contribute new insights to the literature and enhance our understanding of the subject matter.

Graduation Committee members:
First supervisor: Dr. Igors Skute;
Second supervisor: Dr. Maximilian Goethner

Keywords
Emerging technology, patent analysis, automotive industry, R&D, technology, portfolio management, innovation, text mining

1. INTRODUCTION

Considering that our society is in a period of profound change and rapid transformation, modern Research and Development (R&D) management is continuously confronted with unprecedented increases in complexity, speed, and uncertainty (Peschl, 2022). Accordingly, high-tech firms aiming to improve their competitiveness are obligated to consistently improve their ability to respond to these changes in their environment (Zhou et al., 2021), and to develop compatible strategies to ensure and enhance their organization's survivability (Peschl, 2022). In order to accomplish this, an effective planning of R&D to have the proper strategic allocation of all available resources is needed (Rezaeian et al., 2017). Furthermore, "considering that R&D intensive projects tend to take a long time between innovation and commercialization" (Sun et al., 2023), pursuing the wrong options can lead to missing out on other more promising opportunities. Therefore, the identification of the right R&D opportunities is a challenging task considering it requires a detailed knowledge of past findings and current trends, and enough expertise and deep understanding of emerging technology pathways (Rezaeian et al., 2017). This challenge in combination with the increase of environmental uncertainty, in the literature also known as "*VUCA*" (Volatile, Uncertain, Complex and Ambiguous) (Solheim et al., 2022), makes it "quite common for managers in high-tech organizations to mismanage their R&D projects" (Chandrasekaran et al., 2015).

According to Schoemaker et al. (2018) companies in a VUCA environment require dynamic capabilities such as sensing external change, seizing new opportunities, and transforming organizations in order to enhance a firm's long-term performance. Likewise, the company that would master these three capabilities "will see emerging trends and undercurrents sooner than rivals, position themselves more favourably for future scenarios, and respond more rapidly as the future unfolds" (Schoemaker et al., 2018), which could lead to major competitive advantages. However, regarding these dynamic capabilities that companies should possess, Pelsch (2022) notes that "quite the opposite can be observed when one looks at how most of today's organizations are struggling with our high-speed VUCA environment." Moreover, Pelsch (2022) elaborates on this statement with that in many cases organizations operate in a rather reactive, standardized and repetitive mode rather than with organizational learning and innovation they claim to utilize. Accordingly, this leads to incremental innovations that are only a response or reaction to changes in the market, in technologies, or user needs. The COVID-19 pandemic is a prime example that many current organizations lack the capabilities to respond to a VUCA environment, where first organizations were unprepared and unclear in their responses (Solheim et al., 2022). They secondly did not have the capabilities to respond to the changes COVID-19 brought with them for businesses (Solheim et al., 2022). Therefore, with the facing challenges of a VUCA environment, companies are forced to become more agile, quick, and dynamic with their technological developments. With organizations being bound to their limited availability of resources to explore new business opportunities, being able to detect and identify novel technology patterns before the trends are formed will help to avoid sunk costs made during development, offer better protection against tech lock-ins by competitors, and improve the brand recognition and value. All leading to significant increases in the competitive advantages of a firm.

A possible appropriate indication of R&D competences are patents, which is an important information source used often in science, technology, or business (Wang et al., 2014). According to Wang et al. (2014) "the ability of using patent information to measure R&D activities is attributed to some unique characteristics of patents, such as availability, relatively objectivity and prompt recognition of technological changes". Moreover, according to Jiang et al. (2023) patent statistical features, such as the number of patents, citations, claims, and family size, have an impact on the financial performance of a company. This contributes to patent data being an important source of competitive intelligence, which can be used to gain competitive advantages. Furthermore, with the analysis of patent documents, it is possible to map industry R&D information, review technology trends, research hot spots, and obtain other relevant information regarding R&D (Wang et al., 2014). Even though patents can provide valuable insights into R&D, capturing the emerging trends on the industry level presents a challenge that calls for a need to evaluate signals form the entire industry. This, in turn, calls for an automated approach with the ability to interpret these signals.

The goal of this research is to construct a reproducible approach to assess the technology composition of an industry by analysing the industry's patent portfolios. The ability to assess the technology composition of the patent portfolio is important for a company to benchmark its own position to that of competitors and the rest of the industry. Additionally, funding agencies or individual investors should, among other things, use the assessment of the compositions of firms to determine on which firms possesses a stronger competitive position, and hence could be a better option to invest in compared to competitors.

For this research, the automotive industry is chosen as a case. Due to this study having a focus on technological developments in a VUCA environment, an industry should be chosen that possesses both of these characteristics. In the last few years, the automotive industry has undergone many significant changes caused by regulatory change, partially due to global natural environment concerns, economic downturn, and globalization (Ettlie et al., 2023). Additionally, the ever-rising external pressure of end-users on firms in the automotive industry to become more sustainable can be a challenging task for firms to meet with their current competences. Moreover, despite the benefits of investing in sustainability, there is no guarantee that this will also lead to enhanced economic performance (Sun et al., 2023), restating the importance of the ability to identify weak technological signals before the trend is formed to achieve competitive advantages.

For these reasons the automotive industry is chosen as a case in this research, and the research question in this study reads as follow: "*How can the technology composition of firms in the automotive industry be assessed based upon their patent portfolio?*"

To find an answer to this research question unsupervised text mining techniques will be used on high-tech venture statements. This approach is chosen since the volume of data available for firms regarding new technological opportunities has grown considerably (Kayser & Blind, 2017). Due to this increase in size, number, variety, and complexity of databases in many fields of innovation, a need for new methods of knowledge management is created (Bengisu & Nekhili, 2006), and firms are forced to become more agile, react faster, and becoming more efficient while continuously monitoring developments in their environment. Here is where text-mining, which "structures and aggregates data in a largely automated manner" (Kayser & Blind, 2017), will be of used with the goal "to extract valuable knowledge and information from the textual content of the scientific papers for technology trend analysis" (Li et al., 2019). With the use of automated text analysis technology patterns can be detected before the trend is formed, in a manner that is faster and more efficient than when done manually. Furthermore, in this research it is chosen to analyse the patent titles of firms in the automotive industry due to its availability, relatively objectivity, and prompt recognition of technological changes in order to analyse technology trends and competitive intelligence (Wang et al., 2014). While this paper builds upon already existing techniques and frameworks used in other studies, it aims to create a new insight of detecting technology trends early on by analysing the automotive industry as a case study. Furthermore, by combining the established methods it is attempted to achieve an improved approach on how to assess a firms technology composition based on their patent portfolio. The study will further provide a practical insight into the composition of the R&D portfolios of firms in the automotive industry, while building the approach on already existing debates on strategic R&D management. This is done in pursuit of establishing a better understanding of the technology compositions of firms for future research.

2. THEORETICAL FRAMEWORK & HYPOTHESES/PROPOSITIONS

2.1 Emerging technology development & R&D management

The complex and rapidly evolving environment of companies in the VUCA world has given rise to emerging technologies with transformative and disruptive characteristics, which has changed the landscape of existing industries while creating new economic opportunities and changing the way people live their lives (Ebadi et al., 2022). The emergence and development of emerging technologies provides an opportunity for companies to strengthen their competitive advantage, by achieving cutting-edge technological breakthroughs, improving operation efficiency, and nurture emerging markets (Xu et al., 2021; Stratopoulos & Wang, 2022). Therefore, the ability to detect new technology trends is of critical importance for companies, "as it enables them to identify opportunities and risks quickly and react to them accordingly by formulating appropriate research, development, and innovation strategies" (Ebadi et al., 2022). However, identifying these emerging technologies in their early stages remains challenging for companies. Accordingly, companies must put their effort into understanding how technologies emerge from scientific and technological (S&T) information sources like scientific articles and patents, and in order to identify the technologies in a timely and reliable manner it is of the essence to thoroughly examine relevant S&T trends and their associated references (Xu et al., 2021; Ebadi et al., 2022).

Furthermore, the already difficult task of detecting the emerging technologies is accompanied by the finite resources R&D managers possess. This leads to a need to optimize the use of all the resources a company has in order to gain the maximum number of benefits. A possible tool for optimizing resource allocation are portfolio techniques, which are powerful tools that allow R&D projects to be analysed in a systematic manner providing an opportunity for the optimization of a company's long-term growth and profitability (Mikkola, 2001). In the literature there are many different approaches to analysing a R&D portfolio based on for instance the attractiveness of the market, growth rate, and competitiveness of the market (Mikkola, 2001). Another approach is focussing on the uncertainties related to the R&D projects, where according to Bistille (2016) the allocation of R&D funds across the portfolio "must simultaneously consider uncertainty from research outcomes and from market acceptance of the resulting technologies." A similar approach to this is the research of MacMillan & McGrath (2002), where they identify five different types of R&D projects scored on their market uncertainty and their technical uncertainty. With market uncertainty the authors mean all possible concerns when it comes to creating sufficient demand, possible (further) regulations, and the response of their competitors. The technological uncertainty revolves around the complexity of technology needed for the new endeavour, and if it is feasible and realistic to create with the firms' competences and resources. On these two dimensions the R&D projects are rated which leads to the framework of MacMillan & McGrath (2002) which can be found in Figure 1 below.

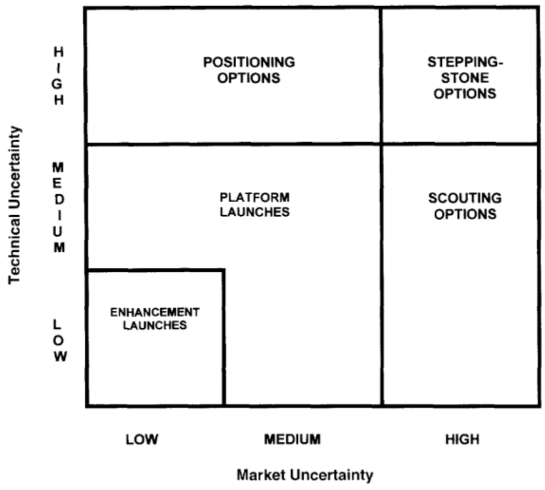

Figure 1. The five types of R&D projects depending on the degree of technical and market uncertainty (MacMillan & McGrath, 2002).

With the help of the framework in Figure 1 the authors try to determine the right category for individual R&D projects, but also enable designing and managing a portfolio of projects that is consistent with a firm's technology strategy. For instance, in the automotive industry a company like Tesla which is known for its strategy around innovative electric cars might need to take a riskier approach in their R&D management in order to maintain their competitive position and keep developing new technologies. Furthermore, MacMillan & McGrath conclude that there is no one correct way to fill in the R&D portfolio, but firms in the automotive industry should rather let their strategy, available resources, and their strategic environment decide on how much to emphasize each category.

2.2 Text-mining

With the current advancements in information and computer technology, the body of public technical literature including scientific papers and patents has grown and starts to be more and more recognized in other research to be an informative source to analyse and study technology trends (Li et al., 2019). A solution to analysing this large body of public technical literature is text mining, which is defined as "the non-trivial extraction of implicit, previously unknown, and potentially useful information from (large amount of) textual data" (Waegel, 2006 as cited in Lee et al., 2023). According to Lee et al. (2023) "text-mining techniques, such as text clustering, text classification, document summarization, latent corpus analysis, and information retrieval, are typically used." Along with the same authors, the basic process of text mining starts with pre-processing, including text clean-up, tokenization, normalization, and stemming.

Many different studies applying text-mining have already been done within the field of new technology creation to analyse large sets of data. An example is the research done by Li et al. (2019), where clustering topics with the Lingo algorithm is applied in order to forecast technology trends. Lee et al. (2023) also uses a Word Frequency Analysis, Association Rule Mining, and topic modelling analysis by performing a latent Dirichlet allocation (LDA) to investigate the innovation in graphene environment technology. Furthermore, Rezaeian et al. (2017) uses clustering with the help of the K-means algorithm to construct a methodological framework for science foresight with the use of the case of natural ventilation technologies.

2.3 The role of Technology Foresight on R&D management

In a VUCA world of increasing complexity, uncertainty, and pace of innovation it has become more difficult for firms to determine in which areas they should innovate. These rapid changes in the environment significantly complicate planning and management of R&D, making methods that address uncertainties valuable (Westphal et al., 2023). One of the methods that support the decision-making process regarding innovation is technology foresight, which helps organizations in maintaining sufficient flexibility for future developments and unforeseen changes in their environment (Keller & von der Gracht, 2014). The successful adaptation of foresight can lead according to Minghui et al., (2022) to enhancing a firms technology management capabilities, improve technology strategic planning, and optimize technology resources allocation. For instance, a firm with high foresight capabilities are more likely to react to weak signals and change their strategy according to the new market demands (Keller & von der Gracht, 2014).

Furthermore, technology foresight is of significant use for R&D management since the projects should be chosen in the context of the environment that will exist at the time the research and technology is completed (Calof & Smith, 2010). In other words, firms should have the capabilities to successfully predict their future environments in order to know which R&D projects will result them in the strengthening their competitive position. Even though there are many different tools for strategic foresight, according to Westphal et al., (2023) the best-know method for foresight is scenario planning. Creating different scenarios can help guide decision making towards a preferred future by exploring a range of options and looking at unexpected consequences based on thinking beyond current, business-as-usual, patterns (Westphal et al., 2023).

Additionally, technology foresight can also be relevant for the automotive industry in helping to detect future technology patterns. The research of Förster (2015), which was limited to only the German automotive industry, stated that there is a lack of information concerning which production technologies might influence sustainable production for automotive suppliers in the future due to stakeholders such as customers and legislation demanding high degrees of sustainability. This will lead to more uncertainty on what the market will demand from the automotive industry regarding sustainability, which could potentially be mitigated with the help of technology foresight. Therefore, it is to be expected that firms with a higher degree of technology foresight capabilities will be able to better understand and position themselves to future market trends.

2.4 The role of Value creation and capturing on R&D management

The emergence of (digital) technologies that are quickly and radically changing in a VUCA environment creates new opportunities for businesses in their continuously changing environment to generate new value creation and value capturing (Mancuso et al., 2023). The advancements in information technology have enabled new approaches to creating and capturing value that extend beyond organizational boundaries, which allows multiple organizations to collaborate in creating and sharing new knowledge (Haim Faridian, 2023). Being able as a firm to successfully create and capture value will lead for instance to a better brand reputation and image, better understanding of customer needs, process optimization, and the improvement of products and services delivered (Mancuso et al., 2023).

To continue, in the current age of manufacturing a trend can be seen in that firms are innovating their business models by shifting from just selling products to selling outcome-based products, where the manufacturer guarantees and takes responsibility for the performance outcomes of their products and services (Sjödin et al., 2020). The guaranteeing of radically higher performance of your products and services should change the way a firm operates their processes, since they now also take responsibility for the outcomes resulting in a need of new delivery processes (Sjödin et al., 2020). This might lead to an increase in uncertainties for the firm by making their delivery of outcomes more complex but will however align their product and services better with the expectations of their customers.

However, in doing so companies should make sure that their value creation proposal is aligned with the needs and the strategy of their business (Costa & Rezende, 2018). A company like Tesla with the mission statement of “Accelerating the World's Transition to Sustainable Energy” (Tesla, About Us, 2023) should not attempt to create new value with innovations that are not in line with their mission and strategy, like developing a new combustion engine and should instead create value through improving their lithium-ion battery pack for instance.

Value creation and capturing is also of great importance in the automotive industry, because when customers perceive that the vehicle they bought meets or even exceeds their expectations it is more likely that they develop a positive image about the brand and remain loyal for any possible future purchases (Suh & Youjae, 2006). This also leads to competitive advantage since the automotive industry is highly competitive and outdoing your competition in creating this value for your customer market could well be the key differentiator in convincing the customer in doing business with you. Therefore, it is to be believed that firms that excel in the creation and capturing of value are more likely to fulfil customers’ demand and enjoy the (financial) benefits from it.

3. RESEARCH DESIGN

3.1 Data

In this paper, Orbis is used as database for collecting data of the patent titles of companies. Orbis is chosen since it is a powerful data resource on private- and listed companies and offers numerous helpful tools to compare and analyze different companies (Orbis | Compare Private Company Data | Bureau van Dijk, n.d.). Furthermore, one of the key factors of Orbis is that it provides strong insights on the intellectual properties (IP) of companies, of which this research will aim to analyze the patents in the automotive industry.

This study’s focus is the automotive industry, of which the definition relates to the ISIC code 291 in the UN Department of Economic and Social Affairs (United Nations, 2008) and the full classification of the code can be found in Appendix A. The code 291 is chosen to establish a focus more on the manufacturing of motor vehicles rather than the manufacturing of components for the motor vehicles, aiming to achieve better and more specific results in the analysis. While filtering the Orbis database on active companies that belong to the code 291 (– Manufacture of motor vehicles), 1,979 companies were retrieved. To further improve the quality of data, filtering to only public limited companies and companies with the minimum of one patent available in Orbis was done. This led to the retrieval of 304 different companies, which together formed the dataset used in this research. Of this dataset the number of patents published by each found company will differ significantly, with the highest number of patens by a single company having a publication number of 581,457 before pre-processing. For this study it is chosen to limit the analysis only to the granted patents, since these patents "asserts that the invention disclosed in the application is indeed new and inventive over the known prior art" (Trippe, 2015), and can "therefore be taken as a quality indicator for innovation activities" (Trippe, 2015). Furthermore, the analysis in this research is conducted on the top 10 most innovative firms in the automotive industry, which is based on the number of granted patents according to Orbis and can be found in Table 1. To limit the focus of the analysis in this research to the top 10 most innovative firms in the industry is chosen, since these firms represent the standards that are set for the whole industry and offer more than sufficient relevant data required for this study. The data of the 10 selected companies obtained from Orbis are collected and converted individually to an .xlsx file and afterwards combined to one Excel sheet before starting the pre-processing process. The following search was done on the 15th of May 2023

Table 1. top 10 most innovative firms in the automotive industry from Orbis.

Company name	Granted Patents
TOYOTA MOTOR CORPORATION	200,128
NISSAN MOTOR CO., LTD.	82,719
GENERAL MOTORS COMPANY	82,697
VOLKSWAGEN AG	68,286
FORD MOTOR CO	66,727
HYUNDAI MOTOR COMPANY	65,670
STELLANTIS N.V.	44,875
RENAULT	29,914
BAYERISCHE MOTOREN WERKE AG	27,457
MERCEDES-BENZ GROUP AG	23,805

3.2 Pre-processing

The patent data of the 10 selected companies retrieved from Orbis are imported into Rstudio, where the pre-processing is started. To start things off, all patents containing non-Latin characters (e.g., patent names in Arabic or Chinese characters) are removed from the dataset. In addition, since patent titles consisting of only a few words provide very limited information, all patent titles having two or less words are also removed from the dataset. The next step is stemming, where with the use of the package 'SnowballC' in RStudio (Bouchet-Valat, 2022) the titles are converted into lowercase and "the basic idea is to reduce the number of words by introducing a common denominator, called a stem, for words that share a common meaning" (Magerman et al., 2010), to aid in comparison of vocabulary. The final step is done with the 'tidytext' package (Silge & Robinson, 2016), where all stop words are removed, and the text is tokenized into a unigram. The term 'stop words' refers to: "all common words that do not contribute to the distinctive meaning and context of documents" (Magerman et al., 2010), with examples of "the", "a", "an", "but", etc. Finally, some custom stop words are also removed from the dataset based on them being to most frequently common terms in the documents, like "vehicle", "method", "system", and "device". This is due to the fact that these words are so common in most patent titles, that including these words will also not contribute to the final results.

3.3 Data analysis

As for the data analysis, topic modelling is used. Topic modelling methods have become a popular approach for topic analysis, as it can handle synonym and polysemy terms very well, and it is helpful in identifying hidden topics from documents (Li et al., 2019). There are several topic modelling methods developed, with a popular one being the Latent Dirichlet Allocation (LDA) since it "is a useful text mining method for detecting latent topics from large size dataset (Wang et al., 2014). This is confirmed by the research done by El Hachem et al. (2023) which found that "if the number of dimensions increases (the size of vocabulary in the LDA increases), the accuracy does not seem to degrade," resulting into LDA being a good method to analyse the large dataset of all patent titles of the selected companies. The process of LDA is also meaningful in this research because it enables automatic mapping, categorization, and classification as well as being a more objective approach than relying on expert opinions (Lee et al., 2023; Magerman et al., 2010). Since, one of the limitations of LDA being that the number of topics cannot be given automatically (Wang et al., 2014), the optimal number k of topics is defined by using the "ldatuning" package in R before running the LDA (Lee et al., 2023).

When the optimal number k of topics is chosen, the LDA helps defining each topic as a distribution of words retrieved from the initial documents (Maulidiya, 2023). Furthermore, it also provides the distribution of words associated with each topic, which support with the interpretation of the topics (El Hachem et al., 2023). The term distribution for each topic is defined by β and the proportions of each topic distribution for the documents is defined by θ, giving us the following two formulas (Grün & Hornik, 2011):

$$\beta \sim \text{Dirichlet}(\delta).$$

$$\theta \sim \text{Dirichlet}(\alpha)$$

The topics are distributed according to a Dirichlet distribution (El Hachem et al., 2023), and there are more complex formulas for calculation the LDA model. However, for this research these formulas are less relevant, and the focus can be more on understanding the basic principles of the LDA model like the two parameters described. With the help of these two parameters, it is possible to better understand the topics created by the LDA. This will help improve analysis in the next part of the research design where the topic labelling takes place.

3.4 Topic labeling

After all patent titles are assigned to a topic generated by the LDA, a manual topic labelling is done. This will be based upon all words distributed to the topic and the common themes and patterns that can be identified for each topic. This can be done by looking at the most frequent words assigned to each of the topics, and investigating if there are any similarities and common themes within each of the topics. Additionally, all the most common words of each topic will be traced back to the patent titles in the original database in order to get a better understanding of what patent titles belong to which topic. Also, the generated topics will be (partially) based upon the current developments in the automotive industry in order to verify their actuality and creating a better understanding of the current R&D projects relevant to this research. In order to conduct the manual labelling in a transparent, non-biased manner the labelling is also done in collaboration with a fellow researcher operating with a similar data set, to assure intercoder reliability. This facilitates a better understanding on the content of each of the topics and what similarities they share. Additionally, it supports a better way to interpret the results of the topics generated by the LDA and improve the results due to the collaboration with a fellow researcher.

4. RESULTS

In this section, the results will be shown of the discussed methods in the research design to the top 10 most innovative firms in the automotive industry based on the data available from Orbis. Firstly, all the patent titles of the individual companies were retrieved to later combine them together, which resulted in a dataset that consisted of 693,686 patent titles. The amount of patent titles dropped down to 356,688 patents after applying the pre-processing steps as described in chapter 3.2. Following, a corpus that is based upon the patent titles was transformed to lowercase letters and removed from all punctuation, numbers, and (custom) stop words and inserted into a document-term matrix. After, with the help of LDA tuning as stated in chapter 3.3, it was possible to determine the optimal number of topics as input for the LDA modelling (Maulidiya, 2023). The package originally provides four different options, but it is standard practice to only look towards two out of the four options. Therefore, in this case CaoJuan2009 and Griffiths2004 were selected. An overview of the metric values determining the optimal number of topics for k can be found in Figure 2 and Table 2, where the former represents a visualization of the LDA tuning results, and the latter provides some of the relevant exact values of these results.

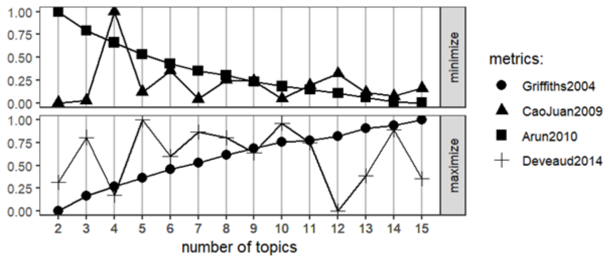

Figure 2. Visualization of LDA tuning results.

Table 2. Metrics LDA tuning.

K	*Griffiths2004*	*CaoJuan2009*	*Arun2010*	*Deveaud2014*
15	-9099256	0.006223869	305477.0	4.332006
14	-9256821	0.004135493	306789.0	4.374933
13	-9326347	0.005199937	310963.8	4.334576
12	-9552233	0.010222010	314889.4	4.303420

The values in the graph and table show that the CaoJuan2009 metric suggests the optimal number of topics is 14, while Griffiths2004 suggests the optimal number of topics is 15. Since there is not one definite answer for the optimal number of topics in this case, some small experiments were done with different numbers for k, ranging from all sorts of values. Based on the given metrics and the investigation of results with different numbers of topics in the end, it is chosen to work with the optimal number of topics set to k = 14 for the remaining of this study. For all of these 14 generated topics the top 20 most frequent terms are computed, which can be found in Appendix B. These will form the basis for the next steps of the analysis, providing an indication about the subjects of each of the topics. For further information, a visualization of the probability of each of these top terms belonging to that topic can be found in the Appendix C, which also gives an early indication to where some of the topics might overlap with each other when it comes to patent titles.

Upon further analysis, it became clear that many topics share a significant amount of similarities in subject with each other. Even though all topics at least have some unique characteristics assigned to them which differentiates them from the rest, it is clear to observe that there are some big common themes among all the topics. These common themes are all relevant to current actual developments in the automotive industry and could provide this study with some interesting and more meaningful results than analysing all the individual topics by themselves. It is therefore chosen that instead of analysing all the 14 topics individually, the more general common themes apparent in these individual topics will be analysed. For this to happen, it is necessary to combine all the individual topics in different groups based upon their common theme based on current industry trends. This is the central theme in the next part of the analysis, where there will be an elaboration on the grounds why each of these topics are combined to each group. Additionally, these groups are labelled with a topic and provide with examples of patent titles that are associated with each of the different groups.

5. TOPICS EXPLAINED

The 14 computed topics are combined into four different groups, which can be found in Table 3 below. This approach is chosen to have a clearer overview of the results and it also the logical step in keeping this research more realistic. The rest of this chapter is focussed on elaborating on these groups and arguments on why they are together.

Table 3. Technology groups created based upon the 14 topics.

Group	Title	from topics
1	Powertrain technology	1, 4, 9, 13
2	Connectivity and autonomous driving	2, 6, 12
3	Electric and hybrid vehicles	3, 7, 8, 11
4	Power and energy optimalization	5, 10, 14

5.1 Powertrain technology

The main focus of group 1 is the powertrain, or in other words an assembly of every component that thrusts your car into motion, like the engine, transmission, driveshaft, and exhaust valves (Kia, 2021). A vehicles powertrain is the entire process from the moment power from the engine is created to when it is delivered to the wheels of the vehicle on the ground (Kia, 2021).

The topics that belong to this group are 1, 4, 9, and 13 of which all share some similarity with each other and the powertrain of the vehicle. However, even though topic 1 holds patents for instance revolving around the engine, wheels and transmission of a vehicle it also includes patents about the overall vehicles body and structure, such as seating, and improving the overall safety and performance of the vehicle's framework. Nonetheless, the majority of the patents in topic 1 revolves around the components that thrust the vehicle in motion. As for the remaining topics 4, 9, and 13, they have a similar focus on the technology of the combustion engine of the vehicles, for instance about the ignition, exhaust gas purification systems, and operating the combustion engine. Despite the similarity the topics differentiate themselves on some small accounts, where topic 4 main focus lies on solely the combustion engine and its transmission and exhaust, topic 9 investigates the technology that looks into the replacement of the combustion engine with electric batteries or lithium batteries. Finally, topic 13 revolves more around the data collection part and improving performance of the powertrain of a vehicle, with examples of technology focussing on detecting systems and methods of operation for combustion engines.

Overall, all topics belonging to group 1 indicate a high connection to the overall components related to the powertrain of a vehicle, therefore the topic of group will be named: *Powertrain technology*. In order to get a better understanding on how this narrative of group 1 is created, some examples of frequent terms belonging to group 1 are displayed in Table 4, which are supported by a potential patent title belonging to this group.

Table 4. Group 1: Powertrain technology, example of some top terms and patent titles.

Top word	Patent title
Combustion	*Self-igniting internal combustion engine*
Exhaust	*How to operate the exhaust gas purification system of an internal combustion engine*
Transmission	*Controller for automatic transmission and method for controlling automatic transmission*
Engine	*Uneven detection device, unevenness detection system, unevenness detection method, data analysis device, and control device of internal combustion engine*

5.2 Connectivity and autonomous driving

Group 2, consistent of topics 2, 6, and 12, revolves around some advance vehicle technologies, that can be divided into two different forms of innovation: connectivity and autonomous driving.

First, topic 2 holds patents that revolve around controlling the vehicles systems and improving the collection and communication of the vehicles data. This includes technology that detects, measures and tracks all data of a vehicle to help improve for instance driver safety, and parking. This collection of vehicular data does not only help the individual driver, but when sharing the data could also benefit the improvement of the entire mobility ecosystem (Insights, 2020). If the collected data is combined with other vehicular data, it will open up an enormous database that will contribute to the evolvement of systems towards “fully autonomous driving, data-driven services, and much greater personalization in the software-defined car” (Ivanov, 2022). As for topics 6 and 12, they hold patents that are aimed at autonomous driving. Where topic 6 is more focussed on safety and assistant technologies while driving, and topic 12 revolves around Advanced Drive Assistance Systems (ADAS). ADAS main goal is to reduce the death and injuries that occur due to car accidents, and to limit danger in car crashes that cannot be avoided.

These two topics are chosen to be combined into one group for this research due to their similarity in nature of uncertainties, and their connection in data management. First, according to research done by StartUs Insights based upon a sample of 4,859 global startups and scaleups concluded that autonomous driving and connectivity are the two most popular new emerging technologies in the automotive industry (Insights, 2020). This reflects the high market demands the technologies have, as they are expected to elevate the automotive industry to the next level based upon their machine learning capabilities. Furthermore, the connection between these two topics enables the sharing and communication of vehicular data, offering autonomous driving a valuable opportunity to subtract new information and enhance its systems. This also goes the other way, the better safety and assistance technologies, the more data that can be collected and communicated further. Therefore, group 2 will be named: *Connectivity and autonomous driving*. Also, for this group some examples of top terms and their patent title counterparts are given in Table 5, in order to create a better understanding of which original data this group is based upon.

Table 5. Group 2: Connectivity and autonomous driving, example of some top terms and patent titles.

Top word	Patent title
Steering	*Systems and methods for recommending a steering angle using visual feedback on a wheel*
Automatic	*Method and system for operating an automatic driving function in a vehicle*
Communication	*Motor vehicle with a communication device, as well as a method for transmitting a data packet*
Data	*Data communication method and data communication system between the service provider and the vehicle*

5.3 Electric and hybrid vehicles

Group 3 holds patents that revolve around electric and hybrid vehicles, more specific around their batteries and other components of the electric and hybrid vehicles powertrain. This group is based upon the topics 3, 7, 8, and 11.

The topics are intertwined with each other through multiple links, so are topic 3 and 9 mostly relevant to electric batteries and electric powertrains of vehicles, while topics 7 and 8 focus more on the hybrid battery and powertrain of vehicles. Another possible way to look at it is by combining the electric and hybrid battery technology (topics 3 and 7) and the electric and hybrid powertrains of the vehicles (topics 8 and 11), showing why there is a case for combing all these 4 topics in this group since they handle at least some similar technologies. The battery technology for both electric and hybrid batteries revolve mainly around systems charging, cooling, and supporting the battery in other ways with devices taking temperature and securing the battery for instance. As for the powertrains of both kinds of vehicles, technologies range from Electric Brake Actuator technology to a shift control device for hybrid electric vehicles.

Overall, all topics belonging to group 3 share some parts of the battery and powertrain of electric and hybrid vehicles and while battery and powertrains are also common themes in group 1 and 4, it is chosen to limit the topic of this group 3 to its uniqueness from the other groups and labelling it with: *Electric and hybrid vehicles*. Furthermore, to create a better understanding of this group, some examples of frequent terms belonging to group 1 are displayed in Table 6, which are supported by a potential patent title belonging to this group. Note that the term 'Electric' is deliberately chosen to be included twice in the table, since it is such a strong theme within this group and helps displaying that a single term can be connected to various different patent titles at the same time.

Table 6. Group 3: Electric and hybrid vehicles, example of some top terms and patent titles.

Top word	Patent title
Electric	*Electric machine-evaporative cooling*
Electric	*Storage element for electric energy and method for producing a storage block*
Hybrid	*Method for operating a serial-parallel hybrid powertrain of a motor vehicle, and motor vehicle*
Battery	*Rechargeable aqueous hybrid battery*

5.4 Power and energy optimalization

Group 4 concerns patents about engine optimization and emission reduction. Even though this group's topics also include some patents regarding the combustion engine and other powertrains mentioned already in group 1. The focus in the current group lays more with optimizing the power and energy systems while reducing emissions, rather than focussing on the whole process of components that are responsible to move the vehicle like in group 1.

Group 4 consists of the topics 5, 10, and 14, of which the first two topics focus more on the efficiency of the engine and topic 14 investigates more into alternative fuelling technologies. In order to optimize the performance of the engine, patents investigate for instance the catalyst used to transform emissions from the engine into steam, or systems that could contribute to emission reduction, and develop methods to prevent misfuelling. The main difference that distinguishes topic 5 from 10 is that topic 5 leans more towards an approach of power and energy systems, while topic 10 is more about engine and emission efficiency.

Overall, this group is a bit more specific in its approach compared to the others and since it main focus is on improving and optimizing already existing processes or investigate some alternatives to already existing components, it is chosen to label this group: *Power and energy optimalization*. An overview of potential top terms belonging to group 4 can be found with the associated patent titles in Table 7.

Table 7. Group 4: Power and energy optimalization, example of some top terms and patent titles

Top word	Patent title
Catalyst	*Method for operating an exhaust gas treatment system having an SCR catalytic converter*
Engine	*Engine lambda dynamic control strategy for exhaust emission reduction*
Combustion	*Method and system for controlling nitrogen oxide emissions from a combustion engine*
Fuel	*Method and system for diagnosing unintended fuelling from fuel injectors of an engine*

6. DISCUSSION AND IMPLICATIONS

In this section, the results of the research are discussed. The goal of the research was to create different topics based on the patent titles retrieved from the top 10 most innovative firms in the automotive industry in order to analyze them. First the discussion will focus on each of the four groups created in the previous section, by analyzing these innovations groups against the current automotive industry developments. Furthermore, these groups will be evaluated on their market and technological uncertainties with the intention to provide implications for their R&D management. Additionally, the role of technology foresight is delved into, along with value creation and capture, assessing their influence on R&D management. Finally, the theoretical implications of this research are discussed, and some insights are given into connecting the R&D portfolios and business strategies of the selected firms with the results of this study.

6.1 Assessment of technology groups

For the firms in the automotive industry dedicating parts of their R&D portfolio to the powertrain technology as described in group 1, it could be of interest to keep an eye out for updates regarding future demand for these types of technologies. Since the automotive industry is already experiencing a monumental shift from internal combustion engine cars to electric vehicles (Valladares Montemayor & Chanda, 2023), companies should be careful not to invest too much into these projects if the future demands are not as certain as before. However, still some potential space for companies to profit from developing more efficient powertrain technologies remains, since with the stricter emission regulation in European and North American countries the demand has increased "for superior powertrains, which are lightweight, and hence, help in increasing fuel economy, decreasing emissions, and improving vehicle performance" (Mordor Intelligence, 2023). On top of that, next to government regulations, the public also keep on shifting their demands for more eco-friendly vehicles, that emphasize on efficiency, which contributes to the growth of the powertrain market (Mordor Intelligence, 2023). This indicates that firms continue to create value for their customers by investing in new powertrain technologies. However, the problem is that in some countries like Norway it will be illegal to sell new petrol or diesel cars in the year 2025, and there are even some speculations that the European Union is thinking of implementing stricter environmental measures around vehicle emissions, which are expected to be very hard to meet with an internal combustion engine (McCann, 2021). Therefore, the market uncertainty of group 1 can be considered medium-high. As for the technical uncertainty, there are not as many technical risks to develop and implement improvements to already existing powertrain technologies. Also, most firms are already investing in these types of technologies, so it does not require that much more infrastructure and capabilities in continuing to improve the already existing powertrain technologies. Due to this the technological uncertainty is perceived as low, which leads to group 1 belonging to the scouting options according to MacMillan & McGrath (2002). The authors advise among other things is that R&D management, when dealing with scouting options, should make plans to scout other markets outside of their current segments. This is where firms with excellent technology foresight capabilities are better in identifying when the perfect moment is for them to exit their current, somewhat outdated, projects and switch to other more demanded technologies like a shift towards lithium batteries for instance.

Compared to the combustion engine and other powertrain technologies, the second group, namely connectivity and autonomous driving technology, is in a much earlier phase in its development. As already mentioned in the previous section connectivity and autonomous driving technology are new emerging technologies in the automotive industry which have a lot of potential. Due to this, the future demand for these complex technologies is relatively high. However, connecting these system technologies could be a bit challenging in directly connecting them towards generating a revenue stream due to these technologies only being a small part of the functioning of the whole vehicle. So, even with a high demand for these types of technologies the commercialization of them could be proven a bit more difficult compared to other technologies. Therefore, the market uncertainty is to be perceived as low-medium. As for developing these technologies, they require a lot of specific skills and resources that are not that easy to possess, for instance large servers, numerous data analyzers, and testing. On top of that, these kinds of techniques have only been recently introduced in the automotive industry and are relevantly complex in nature. Nonetheless, the technical uncertainty is judged as medium capabilities and skills necessary to succeed in developing these technologies. This leads to group 2 belonging to the platform launches according to MacMillan & McGrath (2002). In accordance with them, firms in the automotive industry which enjoy a dominant market position are advised to adopt these projects into their R&D portfolio since for them the technological uncertainties should be manageable and future demands sees very little uncertainty. Even though these complex technologies are an effective way for firms the create and capture value, the firms in the automotive industry should only adopt these technologies into their R&D portfolio if it is in line with their strategy. To elaborate, automotive firms focusing less on being the leading innovator in the industry should maybe invest less into these types of projects since they are still in their earlier stages of development. Rather they should shift their focus to projects more in line with their strategy and needs.

Continuing with the third group, the demand for electric and hybrid vehicles is at an all-time high, with prospects of even more market growth due to upcoming markers for electric vehicles in India, Thailand, and other Southeast Asian countries (Mordor Intelligence, 2023). In order to combat the impacts of climate change the automotive industry is shifting towards mostly electric vehicles as a sustainable transportation solution (Valladares Montemayor & Chanda, 2023). Currently, hybrid and electric vehicles technology are at the center of most R&D portfolios in the automotive industry. However, the technical uncertainty involved with electric vehicles might become a problem later on for the electric industry, since in the manufacturing process of making electric vehicles lithium-ion is of the essence. The problem with lithium-ion is that the electric industry is heavily reliant on the access to this material for its production. With the environmental concerns, scarcity of the material, and human rights concern of people working in a lithium mines the technical future is highly uncertain. Therefore, group 3 should belong to the positioning options. Even though there is a high demand for electric vehicles firms should identify and look for solutions or alternatives for the highly uncertain technological uncertainty. At the moment, it is definitely worthwhile to stay developing electric and hybrid vehicle technologies for the automotive industry. However, automotive firms should utilize their technology foresight capabilities to attempt to mitigate the uncertainty related to lithium-ion batteries and come up with potential future solutions. On another note, according to Valladares Montemayor & Chanda (2023) the current focus on R&D in electric technology is mostly centered around the powertrain (batteries and motors) and has a lack of attention to the impacts of the remaining car components and processes to manufacture them which could potentially be the next innovation step towards a more sustainable vehicle.

Lastly, the fourth group: the technology related to the power and energy optimalization. When practicing the optimalization of the processes in a vehicle it will automatically result in numerous positive benefits, like less fuel being consumed while driving the same distance, or more engine power. All these sorts of benefits are in high demand by customers wanting an eco-friendlier vehicle that is cheaper to operate, which indicated to a low market uncertainty. However, car manufactures need to keep in mind that these benefits only hold true for the customer if the overall price of the vehicle does not increase out of proportion due to these development costs, which could result in less market demand. Therefore, there is still a low-medium market uncertainty for this group. Nonetheless, the technical uncertainty is low in this case, while most companies already possess part of the skills and capabilities required for these technologies and are not as incremental in nature as in other groups. Therefore, group 4 should also belong to the platform launches according to MacMillan & McGrath (2002). As where connectivity and autonomous driving technologies were placed into the same category, they share some similarities in implications for R&D management. However, it should be noted that the technological uncertainty is higher for these complex systems than in the power and energy optimalization. This makes developing these types of technologies of group 4 also more accessible to the smaller companies in the automotive industry that do not necessarily put their focus on being the frontrunners in innovation.

With two of the four topics belonging to the platform launches according to MacMillan & McGrath (2002) framework, and the remaining two belonging to the steppingstone and positioning option. This diversity in results is somewhat predictable since companies in the automotive industry should always diversify their portfolio management. A smart company should never focus on only uncertain technologies, but rather spread their investments in all different kinds of opportunities. The same can be found back in the results of this research, where the ‘less’ riskier options like connectivity and autonomous driving and Power and energy optimalization technologies could help automotive companies protect, and might even expand, their current competitive advantage. Furthermore, the uncertainty in market demand for the Powertrain technology should urgent cautious for firms in not investing too many resources in developing these kinds of technologies that could turn obsolete in a few years. Additionally, the investment in electric and hybrid vehicles technology should be approached with some caution but can nonetheless form a basis for future competitive advantage. Finally, the way firms in the automotive industry should engage in these technologies and diversify their R&D portfolio, relies on their own individual business strategies, needs, and capabilities. Additionally, automotive firms should use foresight to identify weak signals and adapt their strategies accordingly to new market demands like the shift from internal combustion engines to electric vehicles.

6.2 Theoretical Implications

From the outcome of this research several academic implications can be drawn. First, this paper adds to the understanding of detecting technology trends early-on by building on already existing theory. Through combining current methods and literature a new approach is created in this research on how to assess a firms technology composition based on their patent portfolio. With the use of patent analysis and text mining techniques this paper is able to identify emerging technology trends in the automotive industry. As academic research in this field is growing in relevancy, this paper could be further built upon when creating other approaches detecting technology trends. Furthermore, the paper contributes to a better understanding of R&D portfolio management, where there are new insights provided in the composition of R&D portfolios of the automotive industry. Additionally, some further insights are given into how firms should approach the management of these types of technologies in their R&D portfolio.

6.3 Implications for firms in the automotive industry

Due to the way this research has been set up it is difficult to trace each topic back to each of the patent titles and owners belonging to the topic. This unfortunately limits the possibilities to analyze each of the selected firms in the automotive industry. However, it is still possible to judge these firms based upon their market position, mission statement, and their way of creating value for their customers and shortly analyze these strategic business units.

As seen in Table 1 Toyota Motor Corporation is leading by far in granted patents compared to the other companies, indicating their competences in developing new technologies and being potentially the first to enter a new market. When it comes to companies like Volkswagen AG and Ford Motor Co, they hold such a strong position in the automotive market. This leads to that they can allow their R&D portfolio to be a bit more risk averse compared to their competitors, in order to maintain their current competitive advantage. Furthermore, companies like Bayerische Motoren Werke AG (BMW), and Mercedes-Benz Group AG have their emphasis on premium performance and luxury, which results in their portfolio's revolving around innovating their vehicles towards high performance and outstanding customer experiences. It is however interesting to note that in Table 1 both companies have the least amount of granted patents compared to the others, indicating that their innovation strategy is more focused around single developments that are important for their brands. Finally, an examination of Nissan Motor Co., Ltd and General Motors Company is intriguing. These companies, after Toyota Motor Corporation, hold the greatest number of granted patents, solidifying their position in the innovation market. This status empowers them to explore the development of technologies, which may carry higher levels of uncertainty. This means that they will be able to compensate for potential losses in riskier investments by relying on their R&D portfolio diversification, which has many different projects connected to different types of uncertainties.

7. LIMITATIONS AND FUTURE RESEARCH

While this study does contribute to the academic literature, it has some limitations. For instance, the results of this study are based upon the research done in the automotive industry and analyzing another case study will result in different outcomes than the ones presented now. Furthermore, since the shortcoming of expert knowledge about the automotive industry and its trends and developments in this research, the labelling process could have been improved if this was in place. Additionally, this paper used unsupervised machine learning in its approach rather than an in-depth case study. Even though unsupervised machine learning benefited the discovery of unknown patterns, efficiency scalability, and ability to handle complex data in this study, it also has some shortcomings compared to an in-depth case study. By analyzing a case in-depth, there is less reliance on the quality of data and there is more explanation and clarity of the results rather than unsupervised machine learning. Furthermore, an in-depth case study could provide more detailed insights into the technology trends and R&D portfolio management than the unsupervised machine learning provided this study with.

If given more time, similar research could focus on analyzing the whole automotive industry rather than only the top 10 most innovative firms based on their granted patents. This would also include the smaller firms in the industry, which could wield interesting results adding to this research. Furthermore, in the future the analysis could focus on the patent abstract rather than just the title to improve the quality and interpretation of the data source, which could create a better understanding of detecting technology trends in an industry.

8. CONCLUSION

The goal of this research was to construct a reproducible approach to assess the technology composition of a company by analyzing their patent portfolio. The further aim was to extend upon the current academic literature of detecting technology trends early on by analyzing patent titles. In order to realize this goal, the automotive industry was chosen as a case study and the following research question was introduced: *"How can the technology composition of firms in the automotive industry be assessed based upon their patent portfolio?"* In order to answer this research question, this study started off by constructing a theoretical framework with relevant topics and hypotheses that would help understanding the influence of emerging technologies, technology foresight and value creation and capturing on R&D management. Furthermore, the research was designed to help us find common themes in all the retrieved patent titles in order to create topics, which later were group together into 4 different groups representing relevant current developments in the automotive industry. With the analysis of these 4 different groups new insights in the technology trends in the automotive industry were created. This is expanded with some practical implications for firms in the automotive industry based upon their innovation strategies and capabilities and the influence it has on their R&D portfolio management. Moreover, the results of this study helped to create a better understanding of current technologies trends and compositions of R&D portfolios in the automotive industry. By analyzing the patent titles in the automotive industry some novel insights were created that could benefit towards a better understanding of detecting technology trends early on. Finally, in the future, the approach used in this paper could be implemented and improved upon to create a further understanding of detecting technology trends in the automotive industry or be extended towards other industries as well.

9. ACKNOWLEDGEMENTS

To start things off, I would like to thank my first supervisor, Dr. Igor Skute for supporting and aiding me with valuable feedback during the whole process of writing this paper and conducting the research. In addition, I extend my appreciation to my fellow students who were conducting research within the same circle as me. Their assistance and collaboration contributed to the successful completion of this bachelor thesis. Additionally, I would like to express my gratitude to my friends and family for their support throughout the writing process. Their valuable feedback and discussions about the topic played a contributing role in enhancing the quality of this research.

10. REFERENCES

About Tesla—Benefits, Mission Statement, & Photos | JobSage. (n.d.). Retrieved 16 June 2023, from https://www.jobsage.com/companies/about/tesla

Bengisu, M., & Nekhili, R. (2006). Forecasting emerging technologies with the aid of science and technology databases. Technological Forecasting and Social Change, 73(7), 835–844. https://doi.org/10.1016/j.techfore.2005.09.001

Bistline, J. E. (2016). Energy technology R&D portfolio management: Modeling uncertain returns and market diffusion. Applied Energy, 183, 1181–1196. https://doi.org/10.1016/j.apenergy.2016.09.062

Bouchet-Valat, M. (2022). Snowball Stemmers Based on the C 'libstemmer' UTF-8 Library. UTC. https://github.com/nalimilan/R.TeMiS

Calof, J., & Smith, J. (2010). The integrative domain of foresight and competitive intelligence and its impact on R&D management. R&D Management, 40(1), 31–39. https://doi.org/10.1111/j.1467-9310.2009.00579.x

Chandrasekaran, A., Linderman, K., & Schroeder, R. (2015). The Role of Project and Organizational Context in Managing High-tech R&D Projects. Production and Operations Management, 24(4), 560–586. https://doi.org/10.1111/poms.12253

Costa, R. G. G. da, & Rezende, J. F. de C. (2018). Strategic alignment of knowledge management and value creation: Implications on to an oil and gas corporation. RAUSP Management Journal, 53(2), 241–252. https://doi.org/10.1016/j.rauspm.2017.11.001

Ebadi, A., Auger, A., & Gauthier, Y. (2022). Detecting emerging technologies and their evolution using deep learning and weak signal analysis. Journal of Informetrics, 16(4), 101344. https://doi.org/10.1016/j.joi.2022.101344

El Hachem, E.-J., Sokolovska, N., & Soula, H. (2023). Latent dirichlet allocation for double clustering (LDA-DC): Discovering patients phenotypes and cell populations within a single Bayesian framework. BMC Bioinformatics, 24(1), 61. https://doi.org/10.1186/s12859-023-05177-4

Ettlie, J., Muammer, O., & Murthy, R. (2023). R&D Dynamic Capabilities in a Changing Regulatory Context. IEEE Transactions on Engineering Management, 70(1), 98–111. Scopus. https://doi.org/10.1109/TEM.2020.3045650

Förster, B. (2015). Technology foresight for sustainable production in the German automotive supplier industry. Technological Forecasting and Social Change, 92, 237–248. https://doi.org/10.1016/j.techfore.2014.09.010

Grün, B., & Hornik, K. (2011). topicmodels: An R Package for Fitting Topic Models. Journal of Statistical Software, 40, 1–30. https://doi.org/10.18637/jss.v040.i13

Haim Faridian, P. (2023). Leading open innovation: The role of strategic entrepreneurial leadership in orchestration of value creation and capture in GitHub open source communities. Technovation, 119, 102546. https://doi.org/10.1016/j.technovation.2022.102546

Insights, S. (2020, April 29). Top 10 Automotive Industry Trends & Innovations 2023. StartUs Insights. https://www.startus-insights.com/innovators-guide/automotive-industry-trends-10-innovations-that-will-impact-automotive-companies-in-2020-beyond/

Ivanov, I. (2022, June). How Connectivity is Redefining the Automotive Industry. Dotmagazine – Joining the Dots in the Internet Industry. https://www.dotmagazine.online/issues/the-edge/connectivity-redefining-automotive-industry

Jiang, C., Zhou, Y., & Chen, B. (2023). Mining semantic features in patent text for financial distress prediction. Technological Forecasting and Social Change, 190, 122450. https://doi.org/10.1016/j.techfore.2023.122450

Kayser, V., & Blind, K. (2017). 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

Keller, J., & von der Gracht, H. A. (2014). The influence of information and communication technology (ICT) on future foresight processes—Results from a Delphi survey. Technological Forecasting and Social Change, 85, 81–92. https://doi.org/10.1016/j.techfore.2013.07.010

Kia. (2021). What Is a Powertrain? Kia Indonesia. https://www.kia.com/id/discover-kia/ask/what-is-a-powertrain.html

Lee, J. Y., Choi, J. W., Choi, J.-H., & Lee, B.-H. (2023). Text-mining analysis using national R&D project data of South Korea to investigate innovation in graphene environment technology. International Journal of Innovation Studies, 7(1), 87–99. https://doi.org/10.1016/j.ijis.2022.09.005

Li, X., Xie, Q., Daim, T., & Huang, L. (2019). Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology. Technological Forecasting and Social Change, 146, 432–449. https://doi.org/10.1016/j.techfore.2019.01.012

MacMillan, I. C., & McGrath, R. G. (2016). Crafting R&D Project Portfolios: Research-Technology Management: Vol 45, No 5. https://www-tandfonline-com.ezproxy2.utwente.nl/doi/abs/10.1080/08956308.2002.11671522

MacMillan, I., & Gunther McGrath, R. (2002). Crafting R&D Project Portfolios. Research-Technology Management, 45, 48–59.

Magerman, T., Van Looy, B., & Song, X. (2010). Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. Scientometrics, 82(2), 289–306. https://doi.org/10.1007/s11192-009-0046-6

Mancuso, I., Messeni Petruzzelli, A., & Panniello, U. (2023). Innovating agri-food business models after the Covid-19 pandemic: The impact of digital technologies on the value creation and value capture mechanisms. Technological Forecasting and Social Change, 190, 122404. https://doi.org/10.1016/j.techfore.2023.122404

Maulidiya, D. (2023a). Topic Modelling using Latent Dirichlet Allocation (LDA) to Investigate the Latent Topics of Mathematical Creative Thinking Research in Indonesia. 3, 2022. https://doi.org/10.26714/jichi.v3i2.11428

Maulidiya, D. (2023b). Topic Modelling using Latent Dirichlet Allocation (LDA) to Investigate the Latent Topics of Mathematical Creative Thinking Research in Indonesia. 3, 2022. https://doi.org/10.26714/jichi.v3i2.11428

McCann, A.-K. (2021, January). Are lithium-ion batteries ready to completely replace the internal combustion engine? Hitachi High Tech Analytical Science. https://hha.hitachi-hightech.com/en/blogs-events/blogs/2021/01/11/are-lithium-ion-batteries-ready-to-completely-replace-the-internal-combustion-engine/

Mikkola, J. H. (2001). Portfolio management of R&D projects: Implications for innovation management. Technovation, 21(7), 423–435. https://doi.org/10.1016/S0166-4972(00)00062-6

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

Mordor Intelligence. (2023). Automotive Powertrain Systems Market Size & Share Analysis—Industry Research Report—Growth Trends. https://www.mordorintelligence.com/industry-reports/automotive-powertrain-systems-market

Peschl, M. F. (2022). Learning from the future as a novel paradigm for integrating organizational learning and innovation. The Learning Organization, 30(1), 6–22. https://doi.org/10.1108/TLO-01-2021-0018

Rezaeian, M., Montazeri, H., & Loonen, R. C. G. M. (2017). Science foresight using life-cycle analysis, text mining and clustering: A case study on natural ventilation. Technological Forecasting and Social Change, 118, 270–280. https://doi.org/10.1016/j.techfore.2017.02.027

Schoemaker, P. J. H., Heaton, S., & Teece, D. (2018). Innovation, Dynamic Capabilities, and Leadership. California Management Review, 61(1), 15–42. https://doi.org/10.1177/0008125618790246

Silge, J., & Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. The Journal of Open Source Software, 1(3), 37. https://doi.org/10.21105/joss.00037

Sjödin, D., Parida, V., Jovanovic, M., & Visnjic, I. (2020). Value Creation and Value Capture Alignment in Business Model Innovation: A Process View on Outcome-Based Business Models. Journal of Product Innovation Management, 37(2), 158–183. https://doi.org/10.1111/jpim.12516

Solheim, M. C. W., Aadland, T., Eide, A. E., & Haneberg, D. H. (2022). Drivers for agility in times of crisis. European Business Review, 35(1), 57–73. https://doi.org/10.1108/EBR-01-2022-0014

Stratopoulos, T. C., & Wang, V. X. (2022). Estimating the duration of competitive advantage from emerging technology adoption. International Journal of Accounting Information Systems, 47, 100577. https://doi.org/10.1016/j.accinf.2022.100577

Suh, J.-C., & Youjae, Y. (2006). When Brand Attitudes Affect the Customer Satisfaction-Loyalty Relation: The Moderating Role of Product Involvement. Journal of Consumer Psychology, 16(2), 145–155. https://doi.org/10.1207/s15327663jcp1602_5

Sun, W., Zhang, X., & Hazarika, N. (2023). Dilemmas of R&D investment risks and sustainability in the clean-tech economy: Evidence from Nasdaq clean edge index components. International Journal of Green Energy, 20(2), 139–152. Scopus. https://doi.org/10.1080/15435075.2021.2023883

Trippe, A. (2015). Guidelines for Preparing Patent Landscape Reports.

United Nations (Ed.). (2008). International Standard industrial classification of all economic activities (ISIC) (Rev. 4). United Nations.

Valladares Montemayor, H. M., & Chanda, R. H. (2023). Automotive industry’s circularity applications and industry 4.0. Environmental Challenges, 12, 100725. https://doi.org/10.1016/j.envc.2023.100725

Waegel, D. (n.d.). The Development of Text-Mining Tools and Algorithms.

Wang, B., Liu, S., Ding, K., Liu, Z., & Xu, J. (2014). Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology. Scientometrics, 101(1), 685–704. https://doi.org/10.1007/s11192-014-1342-3

Westphal, L. M., Sturtevant, B. R., Reese, G. C., Quigley, K. M., Crabtree, J., Bengston, D. N., Fleischman, F. D., & Plisinski, J. S. (2023). Preparing for an uncertain future: Merging the strategic foresight toolkit with landscape modeling in northeast Minnesota’s forests. Landscape and Urban Planning, 237, 104798. https://doi.org/10.1016/j.landurbplan.2023.104798

Xu, S., Hao, L., Yang, G., Lu, K., & An, X. (2021). A topic models based framework for detecting and forecasting emerging technologies. Technological Forecasting and Social Change, 162, 120366. https://doi.org/10.1016/j.techfore.2020.120366

Zhou, W., Su, D., Yang, J., Tao, D., & Sohn, D. (2021). When do strategic orientations matter to innovation performance of green-tech ventures? The moderating effects of network positions. Journal of Cleaner Production, 279, 123743. https://doi.org/10.1016/j.jclepro.2020.123743

11. APPENDIX

11.1 Appendix A ISIC code 291

From the UN Department of Economic and Social Affairs:

291 Manufacture of motor vehicles

See class 2910.

2910 Manufacture of motor vehicles

This class includes:

- manufacture of passenger cars
- manufacture of commercial vehicles: vans, lorries, over-the-road tractors for semi-trailers etc.
- manufacture of buses, trolley-buses and coaches
- manufacture of motor vehicle engines
- manufacture of chassis fitted with engines
- manufacture of other motor vehicles: snowmobiles, golf carts, amphibious vehicles fire engines, street sweepers, travelling libraries, armoured cars etc.
- concrete-mixer lorries
- ATVs, go-carts and similar including race cars

This class also includes:

- factory rebuilding of motor vehicle engines

This class excludes:

- manufacture of lighting equipment for motor vehicles, see 2740
- manufacture of pistons, piston rings and carburetors, see 2811
- manufacture of agricultural tractors, see 2821
- manufacture of tractors used in construction or mining, see 2824
- manufacture of off-road dumping trucks, see 2824
- manufacture of bodies for motor vehicles, see 2920
- manufacture of electrical parts for motor vehicles, see 2930
- manufacture of parts and accessories for motor vehicles, see 2930
- manufacture of tanks and other military fighting vehicles, see 3040
- maintenance, repair and alteration of motor vehicles, see 4520

11.2 Appendix B: top 20 words for each topic.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14
control	motor	cell	motor	control	control	control	battery	control	internal	engine	program	motor	fuel
engine	control	fuel	fuel	power	transmission	battery	engine	combustion	combustion	drive	methods	apparatus	gas
seat	structure	battery	engine	motor	apparatus	electrolyte	processing	internal	engine	battery	structure	control	cell
structure	communication	manufacturing	combustion	transmission	automatic	hybrid	structure	engine	control	electric	engine	driving	exhaust
motor	seat	assembly	controlling	hybrid	driving	assembly	manufacturing	transmission	exhaust	controlling	computer	fuel	manufacturing
front	controlling	secondary	electric	manufacturing	engine	vehicles	equipment	electrode	structure	assembly	systems	automobile	equipment
controlling	mounting	motor	apparatus	storage	support	variable	information	exhaust	apparatus	control	steering	front	motor
rear	vehicles	apparatus	detection	catalyst	structure	electric	electric	controlling	gas	rear	vehicles	transmission	apparatus
variable	unit	producing	assembly	engine	motor	solid	motor	apparatus	purification	material	detection	manufacturing	cooling
valve	hybrid	using	cell	supply	pressure	engine	controlling	manufacturing	cooling	wheel	power	combustion	internal
transmission	operating	seat	brake	electric	determining	material	detection	automatic	hybrid	manufacturing	control	operating	operating
exhaust	storage	stack	gear	exhaust	air	active	control	battery	film	air	estimation	data	valve
wheel	assembly	transmission	exhaust	fuel	shift	steering	front	air	valve	hybrid	hybrid	thereof	using
assembly	rear	body	transmission	gas	hybrid	using	support	lithium	fuel	detection	wheel	information	layer
internal	comprising	power	support	steering	power	brake	using	structure	cylinder	combination	manufacturing	electric	catalyst
body	battery	electrode	drive	operating	fuel	nonaqueous	power	secondary	vehicles	arrangement	sensor	bumper	display
apparatus	air	separator	manufacturing	assembly	vehicles	connection	body	intake	motor	internal	operating	hybrid	electrode
automobile	wireless	storage	using	material	medium	lithium	driving	vehicles	information	exhaust	valve	vehicles	combustion
belt	management	controlling	machine	vehicles	car	methods	including	including	storage	interior	display	assistance	power
equipped	parking	gas	driving	charging	manufacturing	body	mobile	thereof	injection	combustion	management	heat	material

11.3 Appendix C Visualization of the top terms for each topic