Faculty of Business Studies and Economics

Assessment of Patent Quality

Challenges for Patent Intelligence Stakeholders using Digital Technologies

Cumulative Doctoral Dissertation
by the Doctoral Committee Dr. rer. pol.
of the University of Bremen

submitted by
M. Sc. Valentin J. Schmitt

Date of submission: September 06, 2024
Date of doctoral colloquium: October 18, 2024

First Reviewer: Prof. Dr. Martin G. Moehrle
Second Reviewer: Prof. Dr. Jan Gerken

Executive summary

Assessing patent quality is a challenging task in patent management, especially in patent intelligence, for two reasons: firstly, because of the ambiguity of the definition of quality and secondly, because of the handling of a vast amount of patent data, particularly in relation to the qualitative comparative analyses traditionally performed by patent practitioners. This dissertation addresses both challenges and explores how to assess patent quality in patent intelligence using digital technologies by considering patent intelligence stakeholders and their perspectives. After providing a brief contextual background to this dissertation, five publications are presented, one of which defines patent quality and its economic, legal, and technological dimensions, and four of which introduce new methods for assessing these dimensions. For example, machine learning is used to predict the legal quality of patent applications by modeling statutory novelty according to patent law, and text-mining is used to assess the legal quality of a patent by evaluating the patentability of a patent application through semantic comparison. In this context, the dissertation addresses and discusses the challenges associated with the implementation of such digital technologies for patent quality assessment, e.g. the explainability of machine learning algorithms leading to mistrust or the challenge of heterogeneity in text-mining. Entailing several implications for management and scholarship, this dissertation extends stakeholder theory by applying it to patent quality, and provides a more comprehensive and multidimensional view by integrating different stakeholder interests.

Contents

List of figures

List of tables

List of abbreviations and acronyms

AI	Artificial Intelligence
CNIPA	China National Intellectual Property Administration
Doc2Vec	Document to Vector
DPMA	German Patent and Trade Mark Office
DTM	Dynamic Topic Modeling
EPO	European Patent Office
GPT-4	Generative Pre-Trained Transformer 4
IP	Intellectual Property
JPO	Japan Patent Office
KIPO	Korean Intellectual Property Office
LLM	Large Language Models
R&D	Research and Development
ROC AUC	Receiver operating characteristic area under the curve
SEP	Standard-Essential Patent
USPTO	United States Patent and Trademark Office
WIPO	World Intellectual Property Office

List of publications considered

This cumulative dissertation is based on five publications that are listed in **Table 1**. This table contains the title of the publication, the journal (J) or the conference (C) to which the publication was submitted, the authors and the current status (as of August 2024). Publications ***P1*** and ***P5*** have been submitted to a journal and are undergoing the double-blind peer-review process. Furthermore, publications ***P2*** and ***P4*** have already been published following a double-blind peer review. Publication ***P3*** is based on and complementary to publication ***P2***. It was presented at a conference and published in the associated conference proceedings.

Table 1: List of considered publications

	Title	***Journal (J) / Conference (C)***	***Authors***	***Status***
P1	Disentangling patent quality: Using a large language model for a systematic literature review.	Scientometrics 2024 (J)	Schmitt	Under review (in round 2)
P2	Assessment of patentability by means of semantic patent analysis – A mathematical-logical approach	World Patent Information 2023 (J)	Schmitt, Walter, Schnittker	Published
P3	Detecting patent conflicts by means of computer-based feature analysis (Translated from German)	PatInfo 2023 (C)	Walter, Schmitt	Published
P4	Modeling an indicator for statutory patent novelty	World Patent Information 2024 (J)	Schmitt, Denter	Published
P5	Semantic analysis for assessing the standard-essentiality of patents – Opportunities and challenges	IEEE Transactions on Engineering Management 2024 (J)	Herzberg, Schmitt, Walter	Under review

Source: Author.

1 Introduction

Patents are documents granted by governments, giving inventors exclusive rights and legal protection for their inventions within a specific territory and timeframe (Chitale et al., 2020; Walter and Schnittker, 2016). In 2022, 3,457,400 patent applications were filed with patent offices worldwide, and a global growth of 1.7% was recorded (WIPO, 2023). This means that around 80 percent of all technical knowledge is recorded in these patents and thus publicly accessible (Asche, 2017) and it becomes obvious what potential the transformation of this data into knowledge has for the pursuit of technological and entrepreneurial strategies. Consequently, the effort of extracting knowledge from vast patent datasets has given rise to a domain known as patent intelligence, a dimension of patent management (Moehrle et al., 2017). Patent data not only provides technological insights but also valuable information about competitors in a structured format (Adams et al., 2006; Ernst, 2003).

Using the extracted knowledge to pursue corporate strategies, e.g. by adapting patent strategies consisting of a proprietary, a defensive, and a leveraging strategy (Somaya, 2012), involves assessing patent characteristics that reflect patent quality by evaluating proprietary and third-party patents. However, there are two general challenges in determining the quality of a patent: the ambiguity of the definition of quality and the handling of a vast amount of patent data, particularly concerning the qualitative comparative analyses traditionally carried out by patent practitioners.

Regarding the first challenge, there is still no clear distinction between the terms of quality and value, and numerous definitions exist that are not mutually exclusive. Researchers often use these terms synonymously (Rassenfosse and Jaffe, 2018; Squicciarinii et al., 2023) and the literature shows that indicators of patent value and quality overlap (Rassenfosse and Jaffe, 2018), leading to additional confusion. The assessment of patent quality is complicated by the different perspectives of various disciplines, which suggests the application of stakeholder theory (Parmar et al., 2010). For example, considering the stakeholders of patent management, such as (i) inventors, (ii) patent attorneys, and (iii) strategic management (Denter et al., 2023), different interests and demands emerge that shape aspects of patent quality:

(i) Baldini et al. (2007) highlight three primary motivations for inventors, particularly those associated with institutions such as universities, to pursue patenting. Firstly, patents can support research by attracting additional funding. Secondly, patents promote the exchange of knowledge with other entities, thus stimulating further research. Thirdly, inventors can benefit personally through financial rewards and a boost to their professional reputation, a motivation that also applies to internal inventors in companies. For an inventor, a high-quality patent can be an incentive to do further research.

(ii) A distinction must be made between external and in-house patent attorneys (for this and the following, see Krajec, 2020). Many external patent attorneys are primarily driven by financial incentives. Consequently, they tend to pursue so-called *broad-claim-excuse* tactics and formulate overly broad patent claims, which often face rejection. This results in a lengthy examination prosecution in which the attorney’s earnings can significantly exceed those of the original drafting stage. In contrast, many in-house patent attorneys are primarily interested in a quick patent grant and a relatively narrow scope of protection, aligning their approach with the company's strategic objectives (Squicciarinii et al., 2023). For an in-house patent attorney, a high-quality patent may be one that is granted quickly. For an external patent attorney, a high-quality patent may be one with an excessive scope of protection.

(iii) Strategic management can have multiple interests: It wants the patent to generate economic benefits, e.g. through license agreements or the establishment of a monopoly position in the market (Ribeiro and Shapira, 2020), and to withstand legal challenges and provide effective protection against infringement by third parties, e.g. through a broad scope of protection (Marco et al., 2019). For strategic management, a high-quality patent may generate high license incomes.

In addition to the multidimensional nature of patent quality, handling the flood of patent data poses the second challenge. Conventional methods for patent assessment are no longer suitable when it comes to effectively analyzing this vast amount of data. For example, the assessment of an invention’s patentability is a key aspect when evaluating the legal perspective of patent quality, which is typically conducted by patent attorneys or patent examiners (Hall et al., 2004; Schuett, 2013). This assessment of patentability involves a thorough review of the prior art, which may involve hundreds or even thousands of existing patents, scientific literature, and other publicly available information (Foglia, 2007; Marco et al., 2017; Marttin and Derrien, 2018). Manually analyzing such a massive volume of prior art can be a time-consuming and labor-intensive process, but the average patent application is only examined for about 15 to 20 hours, which often results in it getting invalidated upon examination in court (Farrell and Shapiro, 2008). However, promising solutions for streamlining this process are opened up by advances in digital technologies, such as traditional text-mining or artificial intelligence (AI) (Brynjolfsson and McAfee, 2014; Denter, 2022; Petralia, 2020).

Accordingly, the use of digital technologies, e.g. text-mining or machine learning as part of data analytics and AI (Walter et al., 2022), may be suitable to address the continuous growth of digitalized patent information and to develop new methodologies for patent quality assessment. Therefore, this dissertation aims to answer three research questions:

RQ-1: *How can stakeholder theory be used to obtain a consistent understanding of patent quality?*

RQ-2: *How can an assessment of patent quality be conducted under consideration of the stakeholders involved in patent intelligence?*

RQ-3: *How can digital technologies improve the assessment of patent quality for patent intelligence stakeholders and what challenges arise from their use?*

The remainder of this dissertation is organized as follows. Section 2 describes the contextual background by explaining how digital technologies and patent quality assessment are currently implemented in patent intelligence, identifying and describing the stakeholders involved, and, lastly, discussing the stakeholder perspectives on patent quality within patent intelligence. Section 3 presents the research framework in which the publications of the cumulative dissertation are categorized. Additionally, the publications are presented by addressing their background and motivation, methodology, findings and implications for the dissertation. Section 4 discusses the publications by answering the research questions. Section 5 summarizes the dissertation, identifies research limitations, and highlights the contributions for scholarship and management.

2 Contextual background

The management of patents aims to implement a company's technological and corporate strategy in the best possible way (Walter and Schnittker, 2016) and has developed into an important corporate function (Bader et al., 2012; Conley et al., 2013). In this context, various studies (e.g. Cao and Zhao, 2013; Ernst et al., 2016; Zhao et al., 2017) have shown that rather than the sheer number of patents it is the management of a company's patents that determines innovation performance. According to Gassmann et al. (2021b), patent management aims to generate competitive advantages by optimizing the patent portfolio and encompasses the management of a company's patent portfolio, the generation and protection of its patents, the generation of commercial value from its patents, and the exploration of patents (Bader et al., 2012; Gassmann et al., 2021b).[1]

The latter, together with the continuous increase in digitized patent information, has given rise to an area known as patent intelligence, a dimension of patent management as described by Moehrle et al. (2017). According to Wustmans (2019), patent intelligence refers to *"[…] the discovery, organization, analysis and evaluation of patent information for the systematic development of knowledge and its use in business decisions"* (Wustmans, 2019, p. 25, translated from German), which is mostly realized through digital technologies. Using extracted knowledge for the pursuit of corporate strategies includes the assessment of characteristics reflecting patent quality, which depends on the stakeholders involved and their perspectives and therefore represents a challenging task.

To outline the contextual background of this dissertation, the following sections explore digital technologies and patent quality in patent intelligence, the stakeholders of patent intelligence, and, lastly, the stakeholder perspectives on patent quality in patent intelligence.

2.1 Digital technologies and patent quality in patent intelligence

Digital technologies, such as traditional data analytics or artificial intelligence (AI), offer promising solutions to the continuous growth of digitized patent information. These technologies are increasingly applied in patent intelligence and its elements to analyze patents in terms of their characteristics (Denter, 2022; Walter et al., 2022). According to

[1] *Note:* More recent studies include further tasks, i.e. the company's organizational arrangements for patent management and the corporate philosophy regarding patents (Agostini et al., 2019; Moehrle et al., 2018).

Moehrle et al. (2017) and Moehrle et al. (2018), patent intelligence comprises five elements: acquisition, business segment analysis, evaluation and valuation, information use, and prior art analysis. All of these elements are important for successful patent management (Wustmans, 2019). The element of acquisition involves the identification and obtainment of potential business partners, suppliers, customers, and individuals, such as inventors (Moehrle et al., 2018). The element of business segment analysis comprises the assessment of opportunities and risks pertaining to a specific business segment, the exploration of present and future market dynamics, and the review and, if necessary, an adaption of strategies for the purpose of exploiting a company's technological environment (Moehrle et al., 2018). In the element of evaluation and valuation, the monetary and technological value of invention disclosures, proprietary patents, and relevant third-party patents is determined (Wustmans, 2019). The element information use involves the application of various methods and digital technology tools for effectively accessing and utilizing relevant patent information (Moehrle et al., 2017; Wustmans, 2019). Lastly, the element of prior art analysis involves detecting and analyzing already patented inventions and published knowledge that might get in the way of granting a company's patent (Varma, 2014).

All elements of patent intelligence utilize digital technologies in some form to assess patents and their quality. For instance, in the element of acquisition, text-mining can be used for an inventor profiling approach to identify prospective human resources by particularly innovative patents, as presented by Chung et al. (2021) or Moehrle et al. (2005). However, for this dissertation, the elements of information use and evaluation and valuation are particularly relevant, as they form the basis for addressing the research questions and are the intersection of all publications in this dissertation. Therefore, the following sub-sections provide a detailed description of these elements. For a description of the other elements, see Moehrle et al. (2017) and Moehrle et al. (2018).

2.1.1 Information use

According to Moehrle et al. (2018), in order to exploit the element of information use, a company must address such questions as what data basis is required for patent intelligence, who is involved in patent intelligence, and whether additional resources are needed for patent intelligence. The objective of this element is to obtain quantitative or qualitative data and contextualize it within the company's strategies. There are four domains of digitalization trends in patent information databases and interrogation tools that are relevant to this element, namely AI, cloud computing technology, data analytics, and data management (for this and the following, see Walter et al., 2022). While cloud computing technology and data management provide essential bases for handling large amounts of data, they do not directly offer tools that are suitable for patent analysis. In contrast, the domains of AI and data analytics have a profound impact on patent search and analysis, offer promising solutions for assessing patent quality, and are therefore addressed in detail in this dissertation.

Data analytics provide various applications for extracting knowledge from a vast number of patent documents and find application in assessing patent quality (Schmitt et al., 2023; Wittfoth, 2019b). This includes more conventional patent analyses that primarily rely on text-mining approaches, enabling the effective use of unstructured data, as well as data-mining and predictive analyses (Abbas et al., 2014; Walter et al., 2017; Walter et al., 2022). Of particular relevance in the context of this dissertation is text-mining, which can be "*defined as the process of extracting […] implicit knowledge from textual data*" (Jo, 2019, p. 3). According to Feldman and Sanger (2009), text-mining comprises four main steps: preprocessing tasks, core mining operations, presentation layer components, and refinement technique. For example, the task of preprocessing may involve approaches such as part-of-speech-tagging or parsing (Hotho et al., 2005). One possible application of text-mining is to analyze semantic structures, such as *subject-action-object* or *n-grams*, which can be used for semantic comparison and clustering of patents (e.g. Gerken and Moehrle, 2012; Kim and Yoon, 2021; Moehrle and Gerken, 2012).

AI provides various applications for more advanced patent analyses and includes trends like image recognition, machine learning, machine translation, natural language processing and neural networks (Aristodemou and Tietze, 2018; Walter et al., 2022). "*Artificial intelligence […] is the science and engineering domain concerned with the theory and practice of developing systems that exhibit the characteristics we associate with intelligence in human behavior*" (Tecuci, 2012, p. 168). In this context, intelligence can be defined as the ability to learn from actions in order to gain maximum success in achieving specific goals and solving complex problems (Gretzel, 2011). AI is an umbrella term and thus comprises different types of techniques that are capable of performing tasks which require human intelligence (Banh and Strobel, 2023; Castelvecchi, 2016).

Of particular relevance in the context of this dissertation is machine learning, which is by far the most widely used in academia and industry (WIPO, 2019), finds application in assessing patent quality (Erdogan et al., 2024; Wu et al., 2016), and deals with algorithms that can solve tasks autonomously by processing data through learning (Brynjolfsson and Mitchell, 2017). A distinction can be made between three different learning approaches: supervised, unsupervised, and reinforcement learning (Banh and Strobel, 2023; Kühl et al., 2022). All three machine learning approaches mentioned have one thing in common: they are discriminative, i.e. they learn to distinguish data in order to perform classifications (Banh and Strobel, 2023). In general, the best-performing algorithms are neuronal networks (Du et al., 2019), which are particularly suitable for detecting correlations in large datasets (Janiesch et al., 2021) and widely used in patent literature (Chen and Chang, 2009; Herzberg et al., 2024; Trappey et al., 2006). However, recent advances in deep neural networks (neural networks with more than one layer (Janiesch et al., 2021)) have led to a new type of deep learning technique that has attracted much attention (Banh and Strobel, 2023; The White House, 2022): generative AI. These generative models can generate new output data by analyzing existing data and are primarily known for large language models (LLM) (Brynjolfsson et al., 2023). "*An LLM is trained by learning to predict the next word in a sequence, given what has come before, using a large corpus of text (such as Wikipedia, digitized books, or portions of the Internet)*" (Brynjolfsson et al., 2023, p. 5) Examples of accessible generative AI tools are ChatGPT, Gemini, GitHub Copilot, Midjourney, or DALL-E (Lee and Hsiang, 2020; Yang, 2023).

2.1.2 Evaluation and valuation

Wustmans (2019) states that "*[…] the element [of evaluation and valuation] is used to summarize the capabilities for determining the monetary and technological value of the company's own invention disclosures and patents as well as relevant third-party patents*" (Wustmans, 2019, p. 55, translated from German). Assessing the value of patents is of particular importance as many patents are worth almost nothing (Harhoff et al., 2003) and understanding their true value helps in making informed decisions. According to Moehrle et al. (2018) and Wustmans (2019), in order to exploit this element, a company must address the questions of how high the monetary value of its patent portfolio is, how high the quality of its patent portfolio is, and – from an outward-looking perspective – how high the quality of its competitors' patents is.

As patents are subject to a high degree of information asymmetry – only the parties directly involved are aware of their true value (Lemley and Myhrvold, 2007) – an assessment of patent value on the basis of market prices is not feasible in most cases.

Methods have therefore emerged that estimate the value based on costs or future cash flows (for this and the following see, Gassmann et al., 2021a; Walter and Schnittker, 2016). Cost-oriented methods determine the value of a patent as the amount that would have to be spent in order to achieve an equivalent future benefit. Market price-oriented methods derive the patent value from the value of comparable patents or the realized market prices of comparable patent transactions. Income-oriented methods determine the patent value from expected future cash flows, e.g. through licenses.

Further research has been conducted on estimating patent value according to various information obtained from patent databases, e.g. about renewals and family size (Harhoff et al., 1999; Pakes and Schankerman, 1984), or investigating the influence of patent characteristics on firm value, e.g. citations (Hall et al., 2005; Trajtenberg, 1990). Some sophisticated approaches apply text-mining and AI techniques to assess patent value. For example, researchers Han and Sohn (2015) found that the semantic similarity between a patent and its forward-cited patents has a significant impact on its survival time – an indicator used as a proxy for patent value (Fan et al., 2023). In addition, experts use deep learning models to estimate the value of a patent. For this purpose, various indicators of patent value are determined and used to train deep-learning models (Aristodemou, 2020; Trappey et al., 2021).

However, the element of evaluation and assessment in patent intelligence presents a challenge. As suggested by Wustmans (2019) and Moehrle et al. (2018), companies are required to assess both the monetary value and the technological quality of patents. Yet there is no clear distinction between these terms, as numerous definitions of quality and value exist that are not mutually exclusive. Researchers often use the terms value and quality synonymously (Rassenfosse and Jaffe, 2018; Squicciarinii et al., 2023) and some consider patent value to be pertinent to patent quality, while others regard value as a proxy for quality (Plečnik et al., 2022). For example, although they use the number of citations a patent receives as an indicator to determine patent quality, Dang and Motohashi (2015) quote from the patent literature, showing that "*[...] frequently cited patents have been proven to have higher technological and economic value*" (Dang and Motohashi, 2015, p. 138). Moreover, both the value and the quality of patents are measured in terms of renewals (Harhoff et al., 1999; Higham et al., 2021), family size (Lanjouw, 1998; Lanjouw and Schankerman, 2004a; Pakes and Schankerman, 1984), or validity (Chen and Zhou, 2018; Mann and Underweiser, 2012), underpinning the recent literature which states that indicators of patent quality and patent value often overlap (Rassenfosse and Jaffe, 2018). Therefore, in this dissertation, as by Higham et al. (2021), the term quality is used to capture the concepts of quality, strength, value, etc. in one word.

2.2 Stakeholders of patent intelligence

Patent intelligence encompasses elements that enable an organization to identify, evaluate, and strategically use relevant patent-related information (Moehrle et al., 2017; Walter and Schnittker, 2016). Digital technologies play a central role in this context. The research and analytical results derived from these technologies are of interest to various stakeholders in the field of patent intelligence.

The consideration of stakeholders and different perspectives has already found its way into a number of disciplines, including law, healthcare, environmental policy, ethics and finance (Parmar et al., 2010). In this context, the term stakeholder describes "*any group or individual who can affect or is affected by achieving the organisation's objectives*" (Freeman, 1984, p. 46). In a business context, stakeholder theory addresses "*the problem of value creation and trade*", "*the problem of the ethics of capitalism*" and "*the problem of managerial mindset*" (Parmar et al. 2010, p. 405) and assumes that different stakeholders hold different interests, as each stakeholder has a unique set of expectations, requirements, and values (Freeman, 1984; Ozdemir et al., 2023). *The problem of value creation and trade* addresses the question of how value is created and traded in a global business context that is subject to rapid change (for this and the following, see Parmar et al., 2010). *The problem of the ethics of capitalism* addresses the connections between capitalism and ethics and lastly, the *problem of managerial mindset* is concerned with the question of how managers should think about management to better create value and link business and ethics. Despite its significance in stakeholder theory, there has been little focus on defining and measuring value creation for stakeholders (for this and the following, see Harrison and Wicks, 2013). A significant portion of the stakeholder literature revolves around debates about managerial responsibilities (e.g. Donaldson and Preston, 1995; Freeman, 1994). These debates are essentially concerned with who should rightfully benefit from the company's activities. However, such discussions often assume that the only relevant form of value is of the economic kind (Freeman, 1984; Freeman et al., 2007). While economic returns are crucial, other factors are often important to stakeholders as well (Bosse et al., 2009). "*Attention to these other factors may prove critical to understanding why firms succeed over time [...]*" (Harrison and Wicks, 2013, p. 98). In summary, stakeholder literature highlights the need for a more thorough assessment of stakeholder requirements.

In order to meet the need for a more thorough assessment of stakeholder requirements, the present sub-section describes the respective tasks and techniques of patent intelligence stakeholders (see **Figure 1**). According to Denter et al. (2023), the following stakeholders contribute to or benefit from patent intelligence: data providers, intellectual property (IP) & technology as well as business analysts, patent examiners and attorneys, human resource management, and strategic and product management (structured according to the type of analysis performed and information used). The contributing stakeholders in particular use digital technologies and methods as part of the information use element to effectively access and analyze patent information. In addition, they are responsible for the evaluation and assessment element by evaluating patents in terms of their monetary and technological quality to support stakeholders such as strategic management.

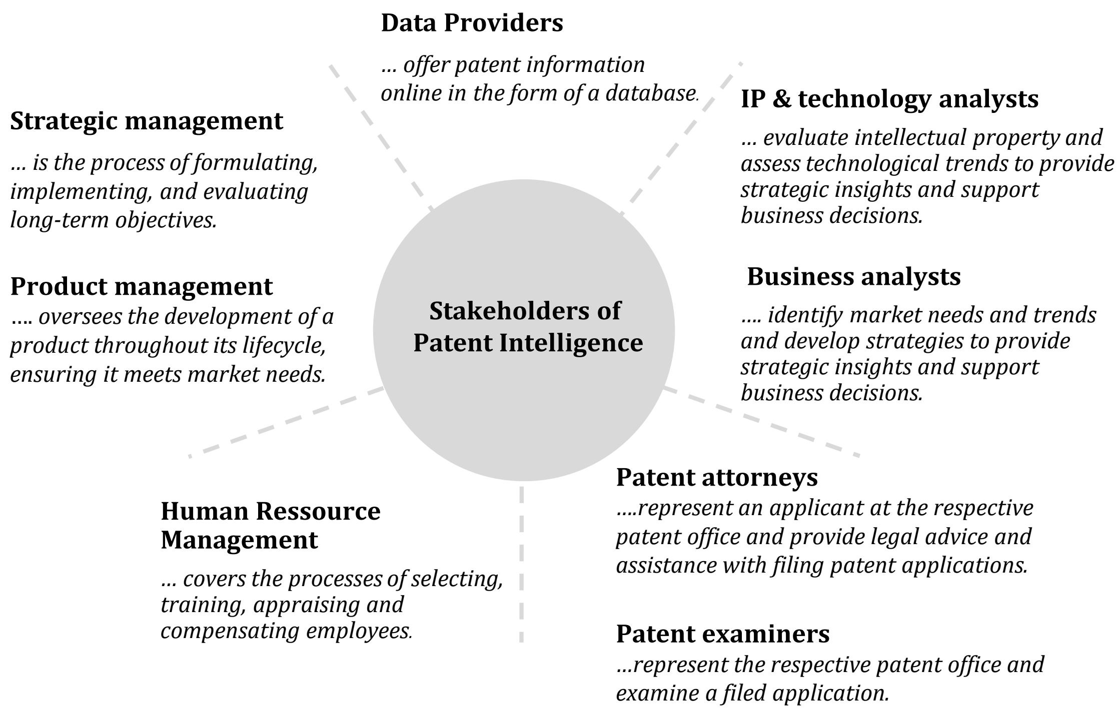

Figure 1: Stakeholders of patent intelligence
Source: Author.

2.2.1 Data providers

The basis for all analyses conducted by stakeholders is created by the data providers, whose contribution to patent intelligence is making patent data available. Data providers offer patent information online in the form of a database (Walter and Schnittker, 2016). A distinction can be made between public and commercial data providers (Walter et al., 2022). The databases of various patent offices, such as the European Patent Office (EPO), China National Intellectual Property Administration (CNIPA), Japanese Patent Office (JPO), Korean Intellectual Property Office (KIPO) and United States Patent and Trademark Office (USPTO) are freely accessible. These patent offices cover approx. 85% of all patent applications filed world-wide in 2022 and thus account for the majority of available and usable technological information (WIPO, 2023). Also, reference is often made to the database of the World Intellectual Property Organization (WIPO), which offers a kind of international patent application under the Patent Treaty Cooperation, granting patent protection in all participating countries (Rassenfosse et al., 2013; Risch and Krestel, 2018, 2019). In addition, these databases integrate various tools that support search and analysis tasks (for this and the following, see Walter et al., 2022). For example, all of the aforementioned databases offer machine translations, and a few, such as the German Patent and Trade Mark Office (DPMA) or the EPO, provide more sophisticated tools based on machine learning. A stronger focus on such sophisticated tools using machine learning or natural language processing can be found among commercial providers of databases and software products. There are providers that sell patent search and analysis solutions, such as Patsnap, which support freedom-to-operate and prior art search with AI tools[2] or LexisNexis, which offer a tool that determines the standard-essentiality of patents[3].

2.2.2 IP & technology analysts and business analysts

IP & technology analysts as well as business analysts research patent data using statistical analysis and technologies in the context of patent management to identify trends and predict business outcomes, thus delivering data-driven recommendations for strategic decisions (Griffiths, 2024; Olavsrud, 2022; White, 2023). They rely on the information made available by data providers and contribute to patent intelligence by analyzing patents and gaining insights into technology fields and business segments. These

[2] Cf. https://www.patsnap.com/solutions/patent-analytics/, last accessed 10 June, 2024
[3] Cf. https://www.lexisnexisip.com/solutions/ip-analytics-and-intelligence/iplytics/, last accessed 10 June, 2024

stakeholders are primarily focused on conducting micro and macro analyses of patents, utilizing metadata- and text-based approaches to extract knowledge from patents.

At the micro level, both stakeholders analyze individual patents in order to assess their characteristics. This involves analyzing their own patents and those of third parties to identify novel (Gerken and Moehrle, 2012; Walter et al., 2017), promising (Park et al., 2013), or valuable patents (Hall et al., 2005). Additionally, this may involve the identification of potentially infringing patents according to semantic similarities (Bergmann et al. 2008). At the macro level, IP & technology analysts as well as business analysts can carry out trend analyses to identify and quantify long-term developments by means of metadata-based approaches (for this and the following see, Walter and Schnittker, 2016). The objective of such analyses is to formulate predictions in order to derive qualitative and quantitative statements about the future development of company-relevant factors such as technology design or market share. For the purpose of formulating statements about trends, patent applications can be examined over time to recognize new technological trends or emerging competitors (Caviggioli, 2016; Niemann, 2015; Xu et al., 2021). For example, the annual reports published by several patent offices can be accessed, which provide information on patent activities in specific technology fields and regions (DPMA, 2022; WIPO, 2022). As an alternative, citation networks can be created. Citation analyses examine the references contained in patent documents and examine their connections (Albert et al., 1991; Narin, 1994). The number of citations can be used to identify inventions that serve as indicators of future technological developments (Harhoff et al., 2003; Trajtenberg, 1990).

In addition to the metadata-based approach, text-based methods, i.e. the aforementioned text-mining and AI techniques can be applied (for this and the following see, Walter and Schnittker, 2016). One solution that is based entirely on text-mining techniques is the patent map. Here, semantic similarities between patents are calculated, which can then be visualized in a two-dimensional representation, i.e. a map, with multidimensional scaling. This method may either be used to identify clusters within a technology field, as patents that are close to each other on the map are semantically similar (Fattori et al., 2003; e.g. Lee et al., 2009; Yoon et al., 2013), or for the general acquisition of knowledge (Cao and Zhao, 2008). A similar method, but one that also takes a temporal component into account, is the patent lane which visualizes *„consistent developments, emerging trends and dormant topics in a technology field"* (Walter and Schnittker, 2016, p. 202 ,translated from German) by calculating similarity values between a patent application and patents from previous years (Block et al., 2021; e.g. Niemann et al., 2017).

Alternatively, topic modeling, a frequently used AI-based technique, can be used to examine a company's technological environment (e.g. Erzurumlu and Pachamanova, 2020; Kaplan and Vakili, 2015; Xu et al., 2021). Topic modeling is an unsupervised learning method used to uncover hidden themes in a collection of texts (for this and the following see, Blei et al., 2003), similar to a patent map. By analyzing the frequency of terms across the documents, topics and their distribution in each document are identified.[4,5] Additionally, machine learning approaches can be used to detect emerging technologies (Lee et al., 2018) or technology opportunities (Lee et al., 2022).

2.2.3 Patent attorney and patent examiner

A patent practitioner or patent attorney represents inventors or companies at the respective patent office, provides legal advice and support in filing patent applications, defends patents, and, if necessary, challenges third-party patents that could affect the client's interests (Walter and Schnittker, 2016). The patent examiner represents the respective patent office and examines the filed application by searching for prior art relevant to the claimed invention that could exclude patentability (Righi and Simcoe, 2019; Walter and Schnittker, 2016). The findings of patent attorneys' and patent examiners' from legal analyses and prior art searches are particularly important and contribute to patent intelligence by providing legal expertise and knowledge.

Depending on the questions addressed by the patent attorney and patent examiner, different types of patent searches and analyses can be conducted (Clarke, 2018). In the following, a detailed description of the patentability, invalidity, and freedom-to-operate search is given.

---

[4] *Note:* A Topic consists of terms that frequently occur together and are semantically related. Essentially, topic modeling helps to organize and understand the content of a text corpus by revealing the underlying thematic structure and provides a way to transform the text of patents into a numeric input for a more sophisticated machine learning model (e.g. Herzberg et al., 2024; Yun and Geum, 2020).

[5] *Note:* To examine patents more effectively from a time-dynamic perspective, dynamic topic modeling (DTM) can be used instead of static topic modeling (e.g. Denter et al., 2019). DTM take into account the fact that topics in a document collection may change over time (for this and the following, see Blei and Lafferty, 2006). Therefore, the patents are grouped by period (e.g. by year) and the topics are assumed to have evolved from the preceding period's set. A static topic model is then created for the first period. Once the topic model of the first period is finished, the model moves on to the next period, creates a new topic model and adjusts it to the previous one. In this way, the algorithm generate a matrix with probabilities of topics for terms that can be considered as a function of time.

A patentability search is a crucial legal analysis that is often performed before filing a patent application (Clarke, 2018; Varma, 2014; Walter and Schnittker, 2016). This search should be carried out by the patent attorney prior to submitting the application and is completed by the responsible patent office examiner as part of the examination procedure. Its purpose is to determine whether the invention in the patent application fulfills the statutory requirements for patentability. For example, the patentability of a claimed invention under U.S. law is defined in the Manual of Patent Examining Procedure (MPEP) based on 35 U.S.C. § 101-103, according to which a claimed invention must be useful, novel, and non-obvious over the relevant prior art (USPTO, 2019a, 2019b, 2019c). To a certain extent, the same applies to patent law as practiced by other patent office's such as the EPO, the KIPO, or the JPO (EPO, 2020a, 2020b; KIPO, 2006; Kowalski et al., 2003). The result of a patentability search should ideally be a comparatively short list of patents and other disclosures that cover the relevant prior art and have been compared with the present patent application and its features (Schmitt et al., 2023).

An invalidity search (also known as a validity search) is typically conducted by an internal or external patent attorney and by a patent examiner if there is a suspicion that a patent application filed by a third-party is not valid (Clarke, 2018; Varma, 2014). Similarly, an opposition search is conducted if an opposition is filed within nine months of the publication of a patent grant, claiming it was wrongfully issued. The objective of both searches is to identify prior art documents that the patent examiner may have overlooked and to demonstrate that the invention in question does not meet the statutory requirements for patentability, e.g. under 35 U.S. C. § 101-103, resulting in the withdrawal of the grant.

A freedom-to-operate search (also known as an infringement search) is a type of search conducted by a patent attorney prior to an industrial activity, such as launching and selling a new product, to determine whether an existing and valid patent could interfere with one's activity and is infringed by one's product (Clarke, 2018; Foglia, 2007; Varma, 2014). Thus, the aim of the freedom-to-operate search is to find out whether there is any geographical and temporal patent protection that could be infringed (Clarke, 2018). It must be ensured that no relevant patent is overlooked during the search (Foglia, 2007). Therefore, it is advisable to carry out a search that aimed at maximum recall (Dirnberger, 2011; Foglia, 2007; Moeller and Moehrle, 2015). Failure to perform this type of search can have costly legal consequences (Fletcher, 1992).

2.2.4 Human resource management

"*Human resource management is a management function that covers the processes of selecting, training, appraising and compensating employees, with respect to regulations in areas including health and safety, labour relations and equal employment opportunity*" (Özbilgin et al., 2014, p. 19). Human resource management is based on the notion that human resources refer to the competencies of employees (Dessler, 2013), which are essential for achieving organizational goals (Özbilgin et al., 2014), and therefore require typical management functions such as planning or controlling (DeCenzo and Robbins, 2005). In conjunction with strategic management, human resource management has evolved and is now more concerned with longer-term, strategic *big-picture* issues and objectives (Dessler, 2013). In this context, Dessler (2013) mentions the tasks of recruitment and placement, training and development, compensation, and finally employee relations.

An employee's knowledge is of particular interest in the tasks of recruitment and placement and training and development, especially as Boutellier et al. (2008) show that research and development (R&D) management is dependent on it. Based on the findings of the contributing stakeholders, a kind of targeted *inventor shopping* can be undertaken in which promising inventors are hired (Chung et al., 2021; Moehrle et al., 2005). At this point, human resource management cannot only consider the individual inventors who may develop new products but must also include their knowledge and acquisition (Imai et al., 2008). In this context, analyses show that by hiring inventors, companies essentially only buy access to the knowledge of the hired inventor. It is accordingly not possible to buy knowledge belonging to the inventor's former employer (Tzabbar et al., 2015).

2.2.5 Strategic management and product management

"*Strategic management can be defined as the art and science of formulating, implementing, and evaluating cross-functional decisions that enable an organization to achieve its objectives*" (David, 2011, p. 6). Here, it is crucial to analyze both the current situation and future developments in order to define objectives and adjust decisions correspondingly (Chon and Olsen, 1990). According to David (2011), the strategic management process comprises three steps: (i) strategy formulation, (ii) strategy implementation, and (iii) strategy evaluation.

(i) Strategy formulation involves several key tasks, including developing a vision and mission, analyzing an organization's external opportunities and threats, assessing internal strengths and weaknesses, defining long-term objectives, and selecting strategies to be pursued (Bourgeois, 1980; for this and the following, see David, 2011; Preble, 1997). In the strategy formulation phase, the organization must make decisions about entering new markets or technologies, allocating resources, expanding or diversifying, entering into international markets, merger or joint venture opportunities, and strategies to avoid takeovers.

(ii) Strategy implementation involves setting annual targets, developing policies, motivating employees, and effectively allocating resources to ensure that the formulated strategies can be pursued (for this and the following, see David, 2011). This step includes fostering a culture that supports the strategy, establishing an appropriate organizational structure, adjusting marketing strategies, creating budgets, developing and using information systems, and aligning employee compensation with organizational performance.

(iii) Lastly, strategy evaluation is intended to determine whether the strategies lead to the desired results (for this and the following, see David, 2011). As both external and internal factors are constantly changing, all strategies are subject to regular adjustments. This involves reviewing the external and internal factors underlying the current strategies, measuring the company's performance, and, if necessary, initiating corrective measures.

The analyses carried out by the contributing stakeholders can serve as input for strategic management in strategy formulation and strategy implementation as well as strategy evaluation. In strategy formulation, for example, the findings of IP & technology analysts can be used to analyze threats to the company, such as emerging competitors and their patents (Hall, 1992), but also internal strengths and weaknesses (Li et al., 2020). In strategy implementation, identified technology trends can be used to focus resources on promising technologies (Li et al., 2020). In strategy evaluation, patents can be used to measure the company's performance (Bloom and van Reenen, 2002) and, if necessary, the overall patent strategy can be adjusted. Thus, the company's competitiveness can be protected by patent fences (Knight, 2013; Lippman and Rumelt, 2003; Somaya, 2012), the commercialization of technologies (Somaya, 2012) can be enabled by opposition proceedings against the granting of third-party patents (Graham S., Hall B., Harhoff D., Mowery D, 2003) or patent rights can be used to generate direct and indirect profits in various contexts (Somaya, 2012).

Product management is often closely intertwined with strategic management. *“Product management is a management concept that is oriented towards the need for cross-functional and cross-divisional control and coordination of products or product groups”* (Aumayr, 2019, p. 7, translated from German). A product manager has extensive market knowledge and is familiar with future market developments and trends, competitor products, or the advantages and disadvantages of products compared to the competition (Aumayr, 2019). This knowledge can be obtained from patent information provided by the contributing stakeholders. Product management can initiate the development of new products based on identified trends or niches (Oh et al., 2020) or the value of a patent (Malewicki and Sivakumar, 2004). Likewise, product management can use the extracted patent information as a stimulus for product development (Aumayr, 2019).

2.3 Stakeholder perspectives on patent quality

All stakeholders involved in patent intelligence consciously or unconsciously evaluate patents in terms of their quality. For example, IP & technology analysts examine whether patents are novel (Gerken and Moehrle, 2012; Walter et al., 2017) or promising (Park et al., 2013), and patent attorneys examine the patentability or validity of a patent (Mann and Underweiser, 2012; Schuett, 2013), as indicators for quality. The assessment of patent quality has already been discussed in this dissertation and the challenges that need to be considered with this concept have been briefly outlined. To further clarify the definition of patent quality in the context of patent intelligence, this sub-section defines quality and its perspectives.

According to ISO 9000, the term quality describes the “*degree to which a set of inherent characteristics of an object fulfills requirements”* (International Trade Center, 2012, p. 7). Suppose one considers the quality of a patent according to said definition. In this case, it immediately becomes clear that every patent has a certain degree of quality if it meets the statutory requirements for patentability, e.g. according to 35 U.S.C. §101-103 under US law. “*Beyond that point, however, [quality] can mean different things to different viewpoints based on different contexts”* (Camarota, 2016, p. 75).

When examining the tasks of patent intelligence stakeholders, it becomes evident that a patent examiner's primary interest lies in said legal quality of a patent, such as patentability. Nevertheless, other stakeholders seem to adopt different perspectives. Schmitt (2024) expands the discussion on patent quality at an abstract level by introducing economic and technological quality alongside legal quality, which includes the aforementioned statutory requirements.

According to Schmitt (2024), a patent has high economic quality when it has the potential to generate financial returns for its owner, influence market dynamics, and create business opportunities. This quality is influenced by external factors such as market demand for the patented technology or product, the competitive landscape, and the patent's role in licensing and collaboration agreements. Internal factors, such as the level of investment in the patent, also play a crucial role.

Moreover, a patent has high technological quality if it has innovative technological attributes and a significant societal impact (for this and the following, see Schmitt, 2024). This includes the originality and innovativeness of the technical solution, the creation of innovation incentives, and the contribution to overall technological advancement. For governments and society, a patent achieves high technological quality if it presents a novel and innovative solution that advances the state-of-the-art and contributes to technological development on a broad societal scale.

Lastly, a patent has high legal quality when it effectively protects innovations and ensures robust legal enforceability for the patent holder (for this and the following, see Schmitt, 2024). This includes a given patentability and validity, a low potential for infringement, and an overall broad scope of protection. The legal quality is significantly influenced by the specific features that define the patent's protective scope.

The three perspectives that determine patent quality – legal, technological, and economic – are considered differently by the stakeholders in patent intelligence. For example, patent attorneys and examiners primarily focus on the legal and technological perspectives in prior art analyses or patent examinations, while all three perspectives are adopted by IP & technology analysts in competitor analysis. **Table 2** outlines the key tasks and interests of patent intelligence stakeholders and indicates their primary perspectives.

Table 2: Perspectives of patent intelligence stakeholders

Stakeholder	*Key Tasks*	*Key Interests*	*Perspective*
Data providers	Data processing, data provision, software applications	Data quality, technological progress	Legal, technological
IP & technology analysts	Trend analysis, patent portfolio analysis, competitor analysis, mergers and acquisitions, patent valuation	Degree of innovation, market potential, economic potential, legal status, technical details, scope of protection, standard-essentiality	Economic, legal, technological
Business analysts	Business segment analysis, trend analysis, market analysis, mergers & acquisitions	Market potential, commercial applicability, technical details	Economic, legal, technological
Patent attorneys	Prior art analyses, patent drafting, patent enforcement	Legal status, scope of protection, patentability, validity, standard-essentiality	Legal, technological
Patent examiners	Patentability examination, patent classification	Prior art, scope of protection, technical details, patentability	Legal, technological
Human resource management	Recruitment, knowledge acquisition	Degree of innovation, technical details, inventor competencies, economic potential	Economic, technological
Strategic and product management	Strategy formulation, strategy implementation, strategy evaluation, product development	Degree of innovation, scope of protection, Market potential, economic potential, standard-essentiality, competitors	Economic, legal, technological

Source: Author.

Furthermore, the key roles of stakeholders highlight that some are directly involved in and perform patent quality assessments, while others primarily benefit from and use these assessments. As an example, one contributing and one benefiting stakeholder and their adopted perspectives are described in more detail below based on the individual stakeholder sections, namely (i) IP & technology analysts and (ii) strategic management, explaining the derivation of the perspectives they adopt.

(i) IP & technology analysts perform trend analyses, evaluate patent portfolios of both their organization and third-party entities, and assess patents for licensing purposes. Consequently, they have a vested interest in understanding the legal characteristics of their patents as well as those of third parties. For example, the scope of protection must to be analyzed to facilitate *invent around* strategies of product management (Reitzig, 2003). At the same time, these analysts seek to identify emerging technology fields, which requires conducting in-depth analyses of the technological details and novelty of patents (Moehrle and Caferoglu, 2019). Additionally, they might assess the economic value of a patent in the context of licensing agreements. Therefore, the quality of a patent is evaluated by IP & technology analysts from the economic, legal and technological perspectives.

(ii) In the strategy formulation phase of strategic management, decision-makers need to assess whether they want to enter a new market or introduce a new technology. To do this effectively, they require legal information about potential competitors and their patents, e.g. in terms of protective scope (Marco et al., 2019). This also includes information on whether there are any freedom-to-operate concerns (Freunek and Bodmer, 2021) or whether licenses for standard-essential patents are required (Herzberg et al., 2024). In the strategy implementation phase, resources must be allocated appropriately. This may mean, for example, investing in the development of inventions within an emerging technology field or in a promising invention (Denter et al., 2022). This allocation of resources necessitates access to technological information. In the strategy evolution phase, the company's performance is evaluated. One method of valuation could be to consider the market value of the company's patent portfolio (Hall and MacGarvie, 2010; Ribeiro and Shapira, 2020). This valuation demands an economic view of the patents that are required. Therefore, strategic management needs information on patent quality from the economic, legal and technological perspectives.

Overall, patent quality in patent intelligence results from a holistic assessment of patents that accounts for the different perspectives and requirements of the various stakeholders. **Table 2** shows how important it is for a company to evaluate patents – whether its own or those of others – not only from a legal but also from an economic and technological perspective, to facilitate decision-making by strategy and product management, human resource management and other stakeholders.

3 Research framework and presentation of publications

This section presents the research framework in which the publications are categorized. On the one hand, this dissertation addresses two key domains of digitalization trends, namely data analytics through the application of text-mining, and artificial intelligence (AI) through the application of machine learning and machine translation. These trends are primarily present in the element of information use in patent intelligence and have a profound impact on patent analyses. On the other hand, this dissertation addresses three perspectives on patent quality – economic, legal, and technological – adopted by patent intelligence stakeholders. Combining the two domains of digitalization trends with the three perspectives on patent quality forms a two-by-three matrix that provides a contextual framework for the publications of this dissertation (see **Figure 2**).

- Publication ***P1*** creates the basis for the subsequent publications by presenting a comprehensive concept for assessing patent quality. It identifies and categorizes patent quality indicators from the relevant literature, aligning them with the economic, legal, and technological quality. These indicators often use data analytics and AI for evaluation. Lastly, the publication provides an agenda for future research on patent quality assessment.
- Publication ***P2*** presents a computer-based process for assessing patentability, i.e. the statutory requirements of novelty and non-obviousness under 35 U.S.C. § 102 and § 103, established indicators of the legal quality of a patent. For this purpose, a mathematical-logical approach and a four-step process are developed to compare patent claims and their features with text-mining techniques.
- Publication ***P3*** is based on ***P2*** and presents a computer-based process for identifying patent conflicts, an indicator for legal patent quality. For this purpose, German patents are translated into English by means of machine translation and then qualitatively compared using text-mining on the basis of semantic structures and patent features of independent claims.
- Publication ***P4*** applies machine learning techniques to patent data and presents an indicator for legal patent quality that estimates statutory novelty as practiced by the USPTO under 35 U.S.C. § 102. For this purpose, a three-step methodology is proposed and deep feed-forward neural networks are trained to analyze patents and estimate their statutory novelty.

- Publication ***P5*** presents a computer-based semantic process for assessing standard-essentiality, an indicator of the economic, legal, and technological quality of a patent, and highlights the differences between the language of technical standards and patents as well as the associated challenges. This process involves text-mining techniques to compare technical standards and patents.

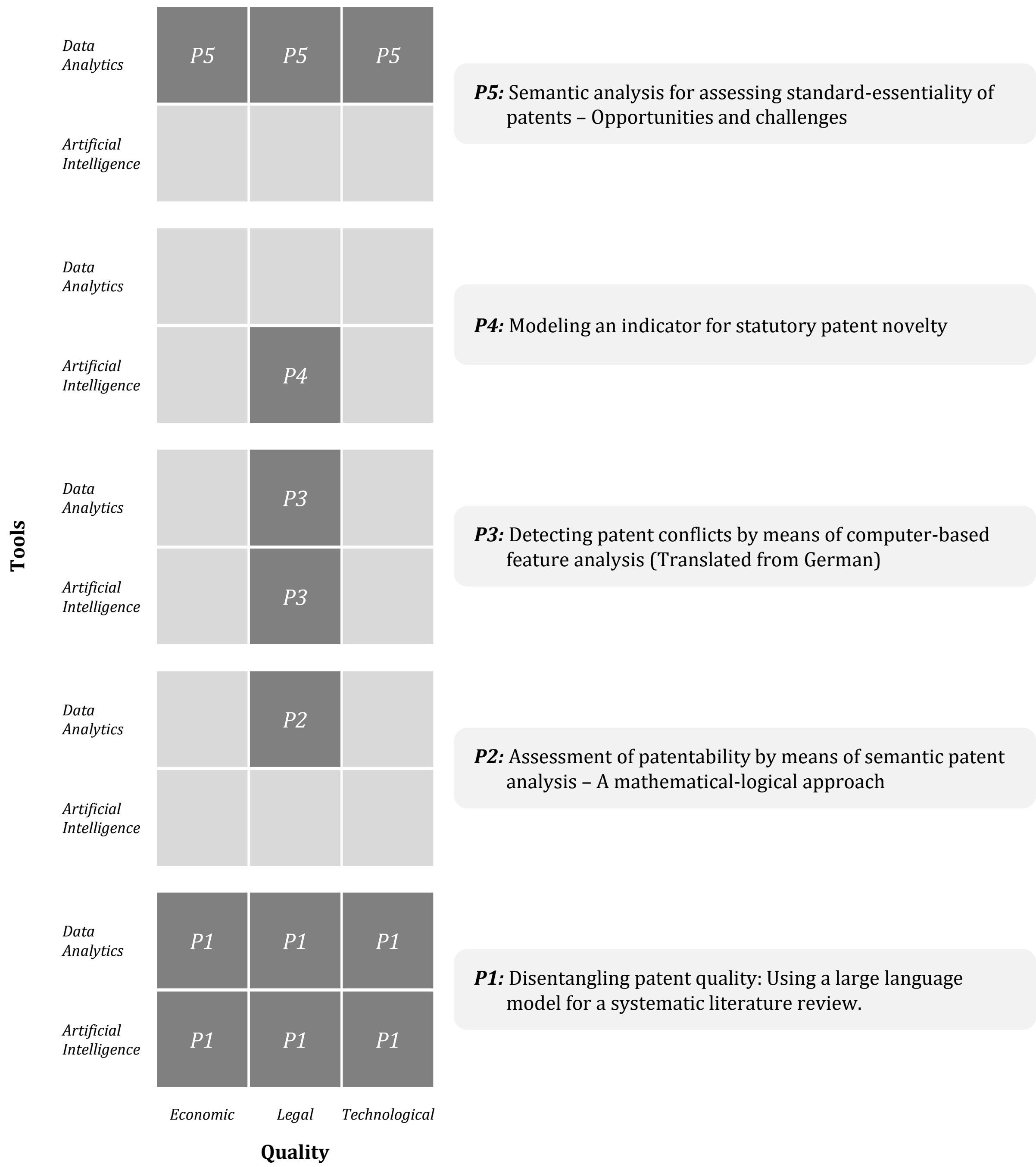

Figure 2: Contextual framework for the dissertation with publications ***P1-P5***.
Source: Author.

The individual publications of this dissertation are summarized in the following subsections. In each case, the background and motivation, the method used and the results of the respective study are presented. Finally, the implications for this dissertation are explained for each publication.

3.1 *P1*: *"Disentangling patent quality: Using a large language model for a systematic literature review"*

Publication ***P1*** presents a systematic literature review, which provides an overview of patent quality assessment in the scientific literature – some using established indicators (e.g. Harhoff et al., 2003), some based on text-mining techniques (e.g. Wittfoth, 2019b) and others based on more complex methods such as AI (e.g. Trappey et al., 2021) – and aligns them with perspectives (here so-called dimensions) of patent quality. ***P1*** was developed under sole authorship, is submitted to the journal *Scientometrics,* and is currently under review.

3.1.1 Background and motivation

The assessment of patent quality has long been a subject in the scientific literature and is becoming increasingly important due to the growing number of patents (Ribeiro and Shapira, 2020; WIPO, 2023). However, as there is no all-encompassing definition, the terms quality and value are often used synonymously, and numerous approaches exist for assessing them (Higham et al., 2021; Plečnik et al., 2022). Recent scientific articles therefore attempt to provide an overview of indicators for assessment and even try to unify the concepts of quality and value (e.g. Ananthraman et al., 2023). These scientific articles have some shortcomings, however, resulting in the absence of a comprehensive overview that combines both concepts and their indicators. Above all, the studies lack a consideration of quality from different perspectives, which is frequently undertaken in the literature and the patent community.

To address these shortcomings, ***P1*** investigates three research questions[6]: (i) *How is the concept of patent quality applied by different stakeholders*, (ii) *what indicators for assessing patent quality exist in the scientific literature and which are suitable for assessing multidimensional patent quality*, and (iii) *how can the assessment of patent quality be*

[6] *Note:* The use of the term research question in section 3 refers to the individual questions addressed in one of the five publications. The abbreviations RQ-1, RQ-2 and RQ-3 describe the overarching questions addressed in the cumulative dissertation.

improved in the future. By answering these research questions, ***P1*** aims to contribute to a more comprehensive understanding of patent quality assessment by incorporating different perspectives and defining dimensions of patent quality.

3.1.2 Methodology

To answer the research questions of ***P1***, a systematic literature review is performed with the support of an LLM. This process involves three steps.

The first step of the process involves the selection of suitable databases in which to search for scientific articles. In the case of ***P1***, the established databases *Scopus*, *Web of Science*, *IEEE Xplore,* and *Google Scholar* are selected. Taking into account the respective taxonomy, search strings are created for each database to obtain a dataset of potentially relevant scientific articles. In the second step of the process, actual relevant scientific articles are identified by means of a screening procedure using both exclusion and inclusion criteria. This involves an automatic screening using the large language model *Generative Pre-Trained Transformer 4* (GPT-4), which evaluates, whether the abstract of an article is relevant to the given research topic on the basis of predefined exclusion and inclusion criteria.[7] Next, a manual full-text screening is conducted to exclude articles that do not specify the assessment of patent quality. In the third step of the process, relevant information, such as patent quality indicators and bibliographic information, is extracted. The indicators are then classified into the three dimensions of patent quality. This classification is performed using GPT-4, which assigns an indicator according to the definitions of patent quality dimensions, followed by a manual validation. Multiple assignments are possible.[8] Furthermore, the relevant scientific articles are assigned to the dimensions of patent quality according to the indicators used and their classification to the dimensions. Finally, a research agenda with propositions is developed based on the results.

3.1.3 Findings

To answer the first research question of ***P1*** (*How is the concept of patent quality applied by different stakeholders?*), terms commonly used in the scientific literature to describe the quality of a patent are combined and stakeholder theory is applied to define

[7] *Note:* To evaluate the selection performance of GPT-4, the inter-rater reliability is determined (Dybå and Dingsøyr, 2008; Kitchenham and Brereton, 2013; Pérez et al., 2020) and other analyses are conducted.
[8] *Note:* To evaluate the categorization performance of GPT-4, the inter-rater reliability is determined (Dybå and Dingsøyr, 2008; Kitchenham and Brereton, 2013; Pérez et al., 2020) and other analyses are conducted.

multidimensional patent quality at an abstract level for the parties involved in the basic patenting mechanism, i.e. patent owner and government. This quality encompasses the three dimensions of economic, legal, and technological quality, and enables the comprehensive assessment of a patent from different perspectives.

To answer the second research question of ***P1*** (*What indicators for assessing patent quality exist in the scientific literature and which are suitable for assessing multidimensional patent quality?*), a systematic literature review was conducted, which identified a total of 762 scientific articles as relevant. Recent literature (Hackl et al., 2023) showing that GPT-4 is a reliable tool and rater for assessing text can be confirmed, by a statistic-significant average Cohen's K of 0.84 between human raters and GPT-4, which indicates an "almost perfect" (K>0.8) agreement (Landis and Koch, 1977). Moreover, a total of 985 different indicators for assessing patent quality were extracted from these articles and assigned to the economic, legal, and technological dimensions and their overlaps, providing an overview of indicators. Once again, the inter-rater reliability is calculated, which shows that the assignment of indicators to the legal and technological dimensions shows a statistically significant "substantial" agreement, while assignments to the economic dimension show a statistically significant "almost perfect" agreement (Landis and Koch, 1977).

A total of 225 indicators are assigned exclusively to the economic dimension. The most common of these are the number of renewals, real options, and stock market reaction to grant events. The most frequent indicators in the legal dimension (168 indicators in total) are the number of legal disputes, grant decisions, and oppositions. The majority of all indicators, namely 276, is assigned to the technological dimension, with the number of forward citations, *International Patent Classifications,* and inventors occurring most frequently. 237 indicators relate to two dimensions, e.g. generality (economic-technological), number of renewals (economic-legal), and number of claims (legal-technological). 36 indicators are assigned to all three dimensions, such as family size, patent age, and *Derwent World Patent Index* family size.

The third research question of ***P1*** (*How can the assessment of patent quality be improved in the future?)* is answered by the formulation of a research agenda with eight research propositions. For example, the assessment of the legal quality of patents is often based on patentability, legal validity, or the number of legal disputes (Farrell and Shapiro, 2008; Mann and Underweiser, 2012; Schuett, 2013), all of which are determined by patent attorneys or examiners and are difficult to assess without legal expertise. Future research could therefore focus on improving the assessment of the legal dimension by introducing new complex indicators or methods that capture or predict the above aspects.

3.1.4 Implications for this dissertation

By defining patent quality as a multidimensional concept and presenting an overview of indicators for determining patent quality, publication ***P1*** contributes to answering ***RQ-1*** of this dissertation. Moreover, in order for patent intelligence stakeholders to effectively assess patent quality, they must first develop a clear understanding of the concept. ***P1*** accomplishes this in several ways and thus contributes to answering ***RQ-2***:

Firstly, a summary of the terms often used in scientific articles to describe patent quality is provided and it is shown how these terms are often used synonymously. Secondly, ***P1*** shows that there is no unambiguous distinction between patent value and quality assessment. Thirdly, ***P1*** defines dimensions of patent quality that can be applied more or less directly to patent quality in the context of patent intelligence and provides an understanding of the three perspectives, i.e. economic, legal, and technological. Fourthly, by identifying the respective indicators for these three dimensions, ***P1*** offers an overview of commonly used indicators. These indicators can be applied by patent intelligence stakeholders to evaluate patents effectively.

Furthermore, by developing a research agenda, ***P1*** lays the foundation for the subsequent publications within the framework of this dissertation. An in-depth analysis of the indicators commonly used in the literature to assess various dimensions of patent quality demonstrates how these assessments can be improved by digital technologies. Many scientific articles, for instance, rely on indicators such as patentability or standard-essentiality to gauge the legal quality of patents (Schuett, 2013; Wittfoth, 2019a), a perspective that must be adopted by the patent intelligence stakeholder and traditionally requires the legal expertise of a patent attorney. Digital technologies offer the potential to develop new methods for determining these indicators, which can reduce the workload of those involved or even replace them in some cases. In particular, the research proposition that the assessment of the legal dimension can be further improved and supported by the introduction of new complex indicators is taken up in publications ***P2*** - ***P5***. Moreover, the in-depth analysis of the patent quality indicators commonly used in the literature reveals a lack of indicators that can be utilized for patent applications, leading to the proposition that new indicators can be developed which enable assessing the three dimensions of patent quality for patent applications. This is addressed by ***P2*** and ***P4*** in which new methods are provided for assessing the statutory patentability requirements of a patent application by means of text-mining and machine learning. ***P5*** also addresses the proposition that additional indicators can be developed to comprehensively capture all dimensions and promote a more holistic assessment by providing new methods for assessing standard-essentiality – an indicator of economic-legal-technological patent quality.

3.2 P2: "Assessment of patentability by means of semantic patent analysis – A mathematical-logical approach"

Publication ***P2*** presents a computer-based process for assessing patentability, i.e. the statutory requirements of novelty and non-obviousness under 35 U.S.C. § 102 and § 103, which are established indicators of legal patent quality. For this purpose, a mathematical-logical approach and a four-step process are developed to qualitatively compare patent claims and their features with text-mining techniques. The method is validated on the basis of an official decision by a USPTO patent examiner. ***P2*** was developed under the co-authorship of Dr. Lothar Walter and Frank Schnittker and published in the journal *World Patent Information*.

3.2.1 Background and motivation

An invention must fulfill statutory requirements in order to obtain a patent grant, as regulated by a national or multilateral body of law. For example, US patent law requires that a patentable invention claims a statutorily patentable subject matter, is novel and non-obvious over the prior art, and useful (USPTO, 2019a, 2019b, 2019c).

The examination of patentability causes an immense amount of manual work, as the most pertinent prior art of an invention must be identified and analyzed. On average, a person skilled in the art needs around 18 hours for this kind of examination (Lemley and Sampat, 2012; Marco et al., 2017), which usually relies on classical methods, such as searches using patent classifications, keywords, or citations (Foglia, 2007; Marttin and Derrien, 2018; Risch et al., 2020). In addition, the examination of patentability involves comparing the patent claims in terms of their subject matter and scope of protection with similar patents in the prior art (Lemley and Sampat, 2012; Marco et al., 2017).

In order to reduce the amount of manual work and resources required in the context of a patentability examination, ***P2*** presents a computer-based four-step semantic analysis process for assessing patentability and examines two research questions: (i) *How can a claim analysis respectively a feature analysis be performed by means of a semantic patent analysis process*, and (ii) *in which way is this kind of process able to support or facilitate statements regarding the patentability of the subject matter disclosed and claimed by a patent application*. By answering these research questions, ***P2*** offers all parties involved in a patent application and the subsequent prosecution the opportunity to assess the scope of protection of claimed inventions, i.e. patent applications, with regard to patentability using text-mining techniques.

3.2.2 Methodology

To answer the research questions of ***P2***, a mathematical-logical approach is presented, in which the features of patent claims are treated as logical statements and the claims as compound and logically connected statements, thus enabling the assessment of novelty under 35 U.S.C. § 102 and non-obviousness under 35 U.S.C. § 103. In this approach, an independent claim of a patent application is considered as features logically linked by AND operations, such as features A, B, and C, reflecting the claimed scope of protection. The claimed combination of features is considered obvious if documents can be found in the prior art which, when considered together, contain the claimed combination of features. If even a single document of the prior art contains the same specific set of features, the claimed combination cannot be regarded as novel, since this document anticipates the invention in its entirety.

Based on the mathematical-logical approach, a semantic process is presented which comprises four steps, namely claim interpretation, prior art determination, similarity measurement, and patentability categorization, whereby the patentability of the claimed feature combination is examined.

Step 1 of the process involves the claim interpretation of a pre-selected patent to interpret the scope of protection in the context of the invention. This includes breaking down the claims into their preambles and features, a semantic and content analysis of the patent's independent claims, and the creation of a language filter. The purpose of this step is to enrich the text of a claim feature with details taken from the description, to provide a more in-depth semantic comparison. In step 2, the prior art of the selected patent is surveyed.

For this purpose, a patent database must be selected, two separate patent search strings for non-obviousness and novelty must be created and the resulting hits extracted. Step 3 involves the similarity measurements between the selected patent and the prior art. The extracted prior art patents are broken down and semantic similarities are calculated for the assessment of non-obviousness and novelty. In step 4, the results of the analysis are evaluated and interpreted with regard to patentability. For this purpose, the calculated similarities are sifted, claim charts are generated and a manual review by a patent attorney is proposed.

As a proof-of-concept, the process is tested using a US-patent application in the field of crane technology and the results are compared to those as determined by a USPTO examiner during office actions.

3.2.3 Findings

To answer the first research question of ***P2*** (*How can a claim analysis respectively a feature analysis be performed by means of a semantic patent analysis process?*), a mathematical-logical approach and a four-step semantic analysis process are presented. Contrary to the common practice of a US-patent attorney or examiner (USPTO, 2019d), the mathematical-logic approach and the four-step semantic analysis process suggest that the non-obviousness requirement should be examined first in the patentability analysis. Using the computer-based semantic process, it is more likely that combined features are found in prior art documents that render the claimed invention obvious, thus leading to a non-obviousness rejection. Unlike other studies that focus on detailed claim analysis and patentability requirements (Alderucci and Ashley, 2020; Ashtor, 2022; Dhulap et al., 2015) or on determining novelty (Chikkamath et al., 2020; Verhoeven et al., 2016), the process specifically addresses the statutory requirements for a patentable invention under 35 U.S.C. § 102 and 103. It offers the advantage of combining theoretical considerations of patentability with classical methods of patent search and semantic similarity to assess both novelty and non-obviousness from a purely statutory perspective.

To answer the second research question of ***P2*** (*In which way is this kind of process able to support or facilitate statements regarding the patentability of the subject matter disclosed and claimed by a patent application?*), the process presented is tested by performing an example of the aforementioned patentability examination for the proof-of-concept. The results indicate that – at least under certain conditions – a semantic process can effectively assess the patentability of a patent application, especially in terms of features related to the prior art, including those disclosed in closely related IP rights. Although the examiner did not directly consider the combination of patents identified by the process, it provides possible alternatives that could have been used to oppose the application.

The computer-based process offers all parties involved in a patent application and the subsequent examination prosecution the opportunity to assess the probability of a successful grant and the scope of protection of the claimed inventions as well as the patentability of the applications. Patent examiners are provided with the opportunity to compare an application with the existing state-of-the-art, focusing on the individual features and combinations thereof in a patent claim. Furthermore, this procedure facilitates the critical review of already granted patents in order to assess their patentability and possible invalidity.

3.2.4 Implications for this dissertation

By presenting a new method that uses digital technologies, in this case text-mining, to determine the legal quality of a patent with the patentability indicator, publication ***P2*** contributes to answering ***RQ-3*** of this dissertation and improves the assessment of legal patent quality in patent intelligence for stakeholders such as patent attorneys.

As ***P2*** points out, determining patentability is a task that involves a high amount of manual work, requiring the analysis of many prior art documents and legal expertise. By using text-mining techniques, the workload of several patent intelligence stakeholders, e.g. the IP & technology analyst or the patent examiner, can be reduced, patent claims can be compared both quantitatively and qualitatively by means of the mathematical-logical approach, and an assessment of patentability can be conducted without a patent attorney or patent examiner. In this way, for example, rejections by the patent office can be avoided and resources can be allocated to other tasks. However, several challenges arise for patent intelligence using text-mining to assess the legal quality indicator of patentability.

The deployment and operation of text-mining are subject to the problem of ambiguity (Fawareh et al., 2008), as this kind of technique is unable to recognize homonyms and synonyms of words, which can lead to noise in the extracted patent information (Arts et al., 2018; Gaikwad, Sonali, Vijay et al., 2014; Helmers et al., 2019). In addition, domain knowledge must be integrated (Akilan, 2015) to create a domain-specific language filter, as done in ***P2***. Moreover, even if patents represent a fairly structured data source, subject to drafting rules (USPTO, 2019e), an inhomogeneous structure (La Justicia de Torre et al., 2018) can exist across all documents, e.g. due to incorrect claim drafting, which can lead to errors when the claims are broken down into features, resulting in information loss. Prior art documents are also subject to heterogeneity (La Justicia de Torre et al., 2018). For a comprehensive analysis of patentability, other disclosures such as scientific articles must be considered in addition to patent documents. Another challenge is to adapt the language and structure used. Therefore, the potentially multilingual (La Justicia de Torre et al., 2018) state-of-the-art has to be taken into account and harmonized. This does not only apply to different document sets but also within one and the same text collection.

3.3 *P3*: *"Detecting patent conflicts by means of computer-based feature analysis"* (translated from German)

Publication ***P3*** introduces a computer-based process for detecting IP rights conflicts, which can be used as an indicator of the legal quality of patents (Crespi et al., 2007). ***P3*** utilizes the same methodology as ***P2*** and compares patent claims and their scope of protection to detect overlaps that could lead to litigation or opposition. The primary advantage of this process is the ability to analyze the multilingual state-of-the-art with the help of machine translation, thus addressing the challenge of multilingualism in text-mining (La Justicia de Torre et al., 2018). The process is validated with a court decision of the Munich patent court. ***P3*** was developed under co-authorship with Dr. Lothar Walter, written in German, presented at the PatInfo 2023 and published in the corresponding conference proceedings. As ***P3*** is methodologically similar to ***P2***, it is only briefly summarized below.

3.3.1 Background and motivation

Even if the subject matter of protection is assessed *ex officio* and examined for patentability, conflicts with IP rights arise in practice, leading to patent litigations and oppositions. In 2021, for example, 841 patent infringement proceedings were pending before German regional courts (Richter and Klos, 2022). To assess a conflict of IP rights, e.g. in the case of patents, feature analyses of independent patent claims, which alone define a patent's scope of protection, are used and compared with the prior art. Such feature analysis is associated with a high level of manual and time-expenditure as the relevant prior art for an invention must be determined via patent classifications, keywords, or patent citations (Foglia, 2007; Marttin and Derrien, 2018; Tseng and Wu, 2008). Also, the claims of the patents must be analyzed and compared with regard to their subject matter and scope of protection (Lemley and Sampat, 2012; Marco et al., 2017).

P3 shows how a computer-based feature analysis can be carried out to identify potential IP rights conflicts using text-mining techniques. Furthermore, ***P3*** deals with the challenge of multilingualism addressed in ***P2*** and demonstrates how machine translation can be used to perform a qualitative feature analysis for non-English language patents.

3.3.2 Methodology

For the identification of potential IP rights conflicts without great manual effort, publication ***P3*** proposes a three-step computer-based process based on text-mining.

Step 1 of this process comprises the preparation of a preselected IP right to translate the IP right – if not available in English – into English and interpreting the scope of protection in the context of the invention. This step is identical to process step 1 in Publication ***P2***, with the difference that the IP right is also translated and utility patents are used additionally. In process step 2, the relevant prior art of the selected IP right is searched. Here too, the step is analogous to step 2 of ***P2***. However, if the IP rights of the searched prior art are not available in English, they must be translated into English for further processing of the claims and descriptions, for example with the help of Google Patents.[9] The search can be carried out in English, as most data providers such as the DEPATISnet[10] from the DPMA provide an English translation of the patents and these can be searched (List, 2012). Lastly, in process step 3, feature comparisons are carried out in order to identify potential overlaps between intellectual property rights with the help of semantic similarity measurements. This step corresponds to step 3 of ***P2***.

As a proof-of-concept, the process is tested using a German patent application in the field of crane technology and the results are compared with those established by the Munich Patent Court.

3.3.3 Findings

Publication ***P3*** introduces a computer-based process that identifies overlapping IP rights in three steps and compares the scope of protection of adjacent IP rights at the feature

[9] Cf. https://patents.google.com/, last accessed 12, June 2024
[10] Cf. https://www.dpma.de/recherche/depatisnet/index.html, last accessed 03, May 2024

level. By using machine translation, the process is able to compare German IP rights, with commonly used English text-mining techniques.

As a proof-of-concept, the process is applied to a selected exemplary DE-patent application. For this purpose, the relevant prior art is searched in the DEPATISnet patent database and a semantic comparison is carried out between the features of the DE-patent application under examination and its prior art. Feature comparisons are performed on the basis of English-language text fragments, whereby the German IP rights are automatically translated using Google Patents. This procedure enables a feature analysis based on multilingual prior art. For example, the five features of the selected DE patent application are semantically compared with over 230,000 text fragments from more than 5,000 German prior art IP rights in the field of crane technology to uncover potential IP rights conflicts. One of the property rights overlaps uncovered by the procedure is the subject of appeal proceedings in which the patentability of the patent application was examined by the patent court in Munich due to an opposition, thus validating the process in relation to non-English IP rights.

3.3.4 Implications for this dissertation

Publication ***P3*** presents a modification of the method introduced in ***P2***. In addition to text-mining, it uses machine translation as a type of AI (Walter et al., 2022) to determine legal patent quality and thus contributes to answering ***RQ-3*** of this dissertation by identifying overlapping IP rights, relevant for patent intelligence stakeholders such as patent attorneys.

As ***P3*** points out, the task of identifying patent conflicts involves a high amount of work, requiring the analysis of many prior art documents and legal expertise. By using text-mining techniques, the workload of several patent intelligence stakeholders, e.g. the IP & technology analyst or the patent attorney, can be reduced. Patent claims can be compared both quantitatively and qualitatively, and claim charts comparing patents and their scope of protection at feature level can be produced without a patent attorney. By using machine translation, this process can also be applied to non-English patents, and yet established text-mining techniques, such as word stemming, developed for the English language, can still be used. Theoretically, patents from the multilingual prior art can thus be compared with each other. However, an additional challenge arises for patent intelligence that uses machine translation as well as text-mining to assess the legal quality indicator of patentability.

The translation of patents can result in noise in the extracted patent information, as the language used in patent claims can differ significantly from ordinary texts (Walter and Schnittker, 2016). "*Claims can be difficult to truly understand for a native speaker, even one with technical understanding of the subject, and an understanding of patents. For a machine [...] this is a major hurdle to be overcome, and human translators will be needed here for some time to come" (List, 2012, p. 194).* For example, terms may be missing that the machine translation model used cannot translate, as is the case in the patent application DE 10 2007 051 539 A1 machine-translated by Google Patents: *"[...] characterized in that the telescopic boom before Austeleskopieren working together with [...]"* (Willim, 2009). The respective patent intelligence stakeholders must bear this in mind, especially when assessing the legal quality of patents, thereby ensuring that reliable assessments of patent conflicts, litigations, or infringements can be made.

3.4 *P4*: *"Modeling an indicator for statutory patent novelty"*

Publication ***P4*** applies machine learning techniques to patent data and presents an indicator for legal patent quality that estimates statutory novelty as practiced by the USPTO under 35 U.S.C. § 102. For this purpose, a three-step methodology is proposed in which, patent data is collected, the required target variable and input variables are constructed and deep feed-forward neural networks are trained. ***P4*** was developed under co-authorship with Dr. Nils Denter and published in the journal *World Patent Information*.

3.4.1 Background and motivation

Novelty in the context of patents is already extensively discussed in scientific articles; however, this concept of novelty differs significantly from that practiced by the patent offices. According to 35 U.S.C. § 102, a patent can only be granted if "*the claimed invention [was] not patented, described in a printed publication, or in public use, offered for sale, or otherwise available to the public before the effective filing date of the claimed invention*" (USPTO, 2019a). So far, novelty has mainly been determined using the theory of recombination, i.e. the novel combination of technology classes (Verhoeven et al., 2016), backward citations (Arts and Fleming, 2018) or word combinations (Arts et al., 2021). Other research measures the novelty of patents based on textual information, e.g. with keywords or document embeddings (Jeon et al., 2022; Lee et al., 2015). None of these approaches derives novelty according to the statutory requirement of 35 U.S.C. § 102, which means that the concept of novelty set forth more or less represents a technological view.

To address these shortcomings, ***P4*** investigates two research questions: (i) *To what extent can the novelty of patents – as measured in the literature – be utilized to assess the novelty of a patent as defined by patent law*, and (ii) *how to model novelty as defined by patent law*. By answering these research questions, ***P4*** aims to contribute to a more comprehensive understanding of patent novelty by raising awareness of misconceptions about how novelty is modeled in economic studies, as opposed to how novelty is practiced by patent offices. Furthermore, ***P4*** presents a new indicator for assessing statutory patent novelty, an indicator for legal patent quality.

3.4.2 Methodology

To answer the first research question of ***P4***, statistical tests are conducted to investigate whether established measures of novelty, specifically those developed by Verhoeven et al. (2016), are suitable for predicting statutory novelty. To this end, logistic regressions are performed using a constructed binary variable for rejection as the dependent variable, and the diverse measures of novelty developed by Verhoeven et al. (2016) as independent variables.

Furthermore, a three-step methodology for modeling a novelty indicator according to USPTO practice is conducted to answer the second research question of ***P4***. The so-called USPTO-novelty-indicator is defined as a measure solely for the statutory requirement of novelty according to 35 U.S.C. § 102, which all patent applications must meet in order to be granted and which is present if a patent application is not rejected for lack of novelty.

In process step 1 of ***P4***, data collection, the dataset is constructed, which comprises patent applications that received no rejection as positive examples of novelty and applications that received a rejection under § 102 as negative examples of novelty. In order to obtain a homogeneous dataset, restrictions are made according to various criteria, for example, filing date and claim length. In step 2, variable construction, the target variable and several input variables are created. The target variable is 0 for all patent applications that were identified as negative patent novelty examples and 1 for all patent applications that were identified as positive patent novelty examples. Furthermore, 353 input variables are created on the basis of linguistic features and patent scope. For this purpose, the patent claims and descriptive texts of each application are collected and preprocessed, e.g. by removing stop words and stemming the remaining words. The claims and description texts are then converted into a vector space using Doc2Vec embeddings with 300 dimensions. Also, various statistical features of the claims, e.g. claim length or complexity, are calculated to determine their depth and width and combined statistical features from patent claims and descriptions are generated. In addition, a variable that reflects the state-of-the-art is created, by counting the number of applications that constitute the pertinent prior art of an application, thereby capturing the retrospective nature of a § 102 examination. In order to account for field-specific and annual-specific differences and behaviors, dummy variables for WIPO technology fields and patent filling years are included as well. In step 3, prediction using deep learning, the classification problem of whether or not a patent application receives a first rejection for lack of novelty under § 102 is solved by training a deep neural feed-forward network model (Kühl et al., 2021). Networks of this type can recognize hidden patterns in the input data and are therefore particularly well suitable for such a task, especially with tabular data (Shrestha et al., 2021). The three-phase approach – data preparation, learning, and evaluation – is used for training and testing deep neural networks (Choudhury et al., 2020; Miric et al., 2022; Shrestha et al., 2021).

The deep learning model is then applied to the entire data set. The result is that each patent application has a metric variable between 0 and 1 which provides information about the probability of potential rejection and indicates the novelty of the patent application as assessed by the USPTO examiners.

The resulting USPTO-novelty-indicator is evaluated in various ways. On the one hand, the indicator is compared with the Verhoeven novelty indicators using a Pearson correlation test, which measures the linear correlation of the said indicators. On the other hand, three confirmatory factor analyses are conducted: one using logistic regression to examine how USPTO novelty affects the likelihood of a § 102 rejection, one converting the novelty variable into deciles to further examine the relationship between novelty and the likelihood of a § 102 rejection by the USPTO, and one comparing the novelty deciles to the actual rejection rates in the dataset. Additionally, the effects of certain input variables on the deep learning model are examined using logistic regression, and finally, the performance of the model in each WIPO domain is evaluated.

3.4.3 Findings

To answer the first research question of ***P4*** (*To what extent can novelty of patents – as measured in the literature – be utilized to assess the novelty of a patent as defined by patent law?)* logistic regressions are conducted, testing whether the various novelty measures developed by Verhoeven et al. (2016) are able to predict a rejection for lack of novelty under 35 U.S.C. § 102. The results show that the different measures of novelty have a minor negative or positive effect on the likelihood of rejection due to lack of novelty. McFadden's pseudo R-square clearly shows that models relying on the novelty measures of Verhoeven et al. (2016) do not contribute to the description of the dependent variable compared to a model that only uses control variables. Consequently, the previously established measures of novelty do not adequately capture the USPTO's definition of novelty and can therefore not be used to assess the novelty of a patent according to patent law.

To answer the second research question of ***P4*** (*How to model novelty as defined by patent law?*), patent novelty is modeled as a quantitative variable based on the practiced definition of the USPTO by following the Define-Operationalize-Confirm process steps (Lambert and Newman, 2022). The resulting novelty indicator is a metric variable between 0 and 1 and reflects the novelty of a patent application according to 35 U.S.C. § 102. In contrast to a binary variable, the indicator has the advantage that it provides information about the probability of a possible rejection.

The evaluation of the USPTO-novelty-indicator provides several insights that can be derived directly from ***P4***. It is noticeable, for instance, that the USPTO-novelty-indicator and the deductively derived novelty indicators point in the same direction to a certain extent, but there are still significant differences between the indicators. Moreover, a higher degree of novelty is associated with a lower likelihood of being rejected under § 102. For example, a patent in the lowest decile of USPTO novelty is almost 60 % more likely to receive a § 102 rejection than a patent in the highest decile. This effect appears to be monotonic, i.e. the higher the decile, the lower the probability. Also, McFadden's pseudo R-squared shows that the probability of receiving a § 102 rejection is much better explained by the modeled USPTO-novelty-indicator than by the already established novelty measures. Furthermore, the number of patent applications rejected under § 102 decreases with increasing deciles of novelty. Finally, ***P4*** shows that the scope of protection has a major influence on the probability that an application will be rejected.

3.4.4 Implications for this dissertation

By presenting a new method using the digital technology of machine learning to determine statutory novelty as an indicator of legal patent quality, publication ***P4*** contributes to answering ***RQ-3*** of this dissertation, improving the assessment of legal patent quality in patent intelligence for stakeholders such as patent examiners.

As outlined in ***P4***, the novelty of a patent has so far mainly been deductively derived from the theory of recombination, thus rather represents a technological concept (Arts et al., 2021; Arts and Fleming, 2018; Verhoeven et al., 2016). However, the novelty determined by the patent intelligence stakeholders patent attorney and examiner is assessed from a legal perspective. The use of machine learning gives all parties involved in the patent examination procedure an initial indication regarding the extent to which a rejection on the grounds of novelty is likely. For example, IP & technology analysts can assess the patentability of an invention before filing the application in order to either obtain the broadest scope of protection or to achieve a grant as quickly and as likely as possible. This represents a clear advantage over established novelty standards, as these are usually based on backward citations that are only available for patent grants. Similarly, the USPTO-novelty-indicator can support examiners faced with a high workload and backlog (Setchi et al., 2021) by being utilized to systematically reject applications for lack of novelty (Choudhury et al., 2020). Another advantage is that the method can be applied to a large amount of data and, compared to the method from ***P2***, can be carried out without any manual work after implementation, which makes it suitable for monitoring tasks. For example, IP & technology analysts and business analysts could use the presented indicator to monitor published applications of potential competitors and strategic management can allocate resources for a novelty search based on the results. However, several challenges arise for patent intelligence using machine learning for assessing the legal quality indicator of novelty.

The performance of machine learning algorithms is highly dependent on the given data quality (Baier et al., 2019; Gregory et al., 2021), and a large dataset is practically a necessity (Boutaba et al., 2018; Brodley et al., 2012). Even if sufficient data is available, other challenges might still arise, for example due to incomplete data, incorrect entries, or biased data, which can lead to inadequate results (Blenk et al., 2017; Kocheturov et al., 2019). Biased data can be a significant problem, especially in the case of patent information resulting from patent examiner decisions. As Kovács (2017) and others have noted, patent examiner decisions are subject to a high degree of bias and noise (Kahneman et al., 2021; Whalen, 2018). Moreover, there is a challenge in terms of the transparency and comprehensibility of machine learning, which are particularly relevant in deep neural networks such as implemented in ***P4*** (Denter, 2022; Leung et al., 2016; Nunes and Jannach, 2017). *"A problem closely related to understanding and transparency is trust. Users will only rely on the results of machine learning models if they genuinely trust them, especially when making important decisions. However, since the level of transparency for many types of machine learning models is still low, trust remains a significant challenge"* (Baier et al., 2019, p. 5). This could be a problem particularly for patent attorneys and examiners, as they may not have the technical knowledge necessary to trust a technique such as the one presented in ***P4***. However, the supervised machine learning report card by Kühl et al. (2021) presented in ***P4*** could generate the necessary trust, as it demonstrates the entire process from model initiation to deployment. Finally, there is the challenge of the complexity of the configuration space, which has already been described by Denter (2022). In addition to the usual hyperparameters, neural networks extend the configuration space with additional ones (for this and the following, see LeCun et al., 2015). When implementing such an algorithm, the number of layers and the number of units per layer have to be chosen. Since neural networks can be prone to overfitting, i.e. the model tends to remember the training data, this choice must be made with care.

3.5 P5: "Semantic analysis for assessing the standard-essentiality of patents – Opportunities and challenges"

Publication ***P5*** presents a computer-based semantic process for assessing standard-essentiality, an indicator of the economic, legal, and technological quality of a patent, and highlights the differences between the linguistic practices of technical standards and patents as well as the associated challenges. By applying the process to 4G technology, it is shown that harmonizing the language of technical standards and patents and accounting for the different structures of both has a positive impact on the assessment of standard-essentiality. ***P5*** was developed under co-authorship with Andre Herzberg and Dr. Lothar Walter, is submitted to the journal *IEEE Transactions on Engineering Management*, and is currently under review.

3.5.1 Background and motivation

When a company is involved in the creation of a specific technical standard or holds owns patents that are likely to be essential to a standard, it is encouraged to declare its patents as standard-essential (Contreras, 2013). However, there is no examination to determine whether the patent is really essential. The declaration is based on the best knowledge and belief of the declaring company. This can lead to an over-declaration, in which a patent is declared as essential to the technical standard without really being necessary for its implementation (van Audenrode et al., 2017) or to an under-declaration, in which a patent that could actually be essential to the standard is not declared as such because the applicant expects economic benefits from this (Brachtendorf et al., 2023; Herzberg et al., 2024). As a result, SEPs are often the subject of litigation (e.g. Colangelo and Aguggia, 2023; Jin and Wang, 2022; Prashant and Ghosh, 2023) involving the determination of standard-essentiality. This determination can be undertaken by various parties such as standardization organizations, patent holders, courts, or independent experts, and is a time-consuming manual process. Numerous standard-essentiality assessments of patents are described in the literature (Brachtendorf et al., 2023; Wittfoth, 2019a). However, the respective approaches do not go into detail about the textual differences between the two elements that need to be compared. Taking into account the different linguistic practices of technical standards and patents, ***P5*** presents a computer-based three-step semantic analysis process for the assessment of standard-essentiality.

The research question examined in ***P5*** is: *How can semantic analysis be used to assess the standard-essentiality of patents, and what are the opportunities and challenges involved*. By answering this question, ***P5*** offers the possibility to assess the standard-essentiality of a patent by means of text-mining and deepens the understanding of the challenges to be considered in such an assessment.

3.5.2 Methodology

To answer the research question of ***P5***, a three-stage semantic analysis process is presented that takes into account the linguistic and structural differences between patents and technical standards, thus addressing the challenge of heterogeneity (La Justicia de Torre et al., 2018). The process comprises the steps of selection (step 1), preparation (step 2), and execution (step 3) and aims to find the best approach to the semantic comparison of a standard and the associated potentially standard-essential patents. The process assumes that truly essential patents have a higher similarity to a technical standard than non-essential patents, as a patent is considered to be standard-essential if one independent claim corresponds to the content of a technical standard (Cho et al., 2023).

Step 1 of the process, selection, comprises the choice of a technology field to be analyzed, the search for the associated technical standard, the selection of a patent database, and the identification and extraction of patents. In step 2, the patents and standards are prepared for the subsequent semantic analysis. For this purpose, a technical standard dictionary is created by compiling all abbreviations contained in the technical standard's abbreviation section and full text. The terminology of the technical standard and the patents is then harmonized and four types of pre-processing are tested, e.g. the abbreviations are fully spelled out or technical terms are abbreviated. Finally, the technical standard and the extracted patents are broken down into different text fragments, e.g. the standard into its sections or the patents into the individual independent claims. In step 3, execution, the semantic analysis is carried out. In order to determine the text fragments and the type of pre-processing that are best suited for assessing the essentiality of a patent, 24 different combinations are considered, e.g. the normative part of the standard and the full claim section of a patent are compared and pre-processed using state-of-the-art techniques. Then, combinations are used to identify similarities based on semantic structures. The resulting similarity values vary significantly and do not provide any information about the best combination because without information on the declaration status, a high mean DSS-Jaccard does not necessarily signify that one combination performs better than another. Using a dataset of truly essential patents from an IP service provider, the combination that generates the highest similarity scores for the truly essential patents is then identified.

Next, the results are discussed and analyzed. For this purpose, the patent that is most similar to the technical standard is compared with the sections of the standard and evaluated based on the assessments of an IP service provider. To confirm the assumption that a high semantic similarity to a standard is associated with a higher probability of being standard-essential, a logistic regression is performed.

3.5.3 Findings

To answer the research question of ***P5*** (*How can semantic analysis be used to assess the standard-essentiality of patents, and what are the opportunities and challenges involved?*), a three-stage process is executed, in which different approaches to comparing potential SEPs with a technical standard are tested and the linguistic and structural differences between patents and standards are harmonized. The results show that advanced text pre-processing, in which technical abbreviations or acronyms are fully spelled out in the processed texts of both technical standards and patents, in combination with individual sections of technical standard and single independent patent claims, leads to the highest mean DSS-Jaccard for true SEPs. As a proof-of-concept, the obtained results are compared with those of an IP service provider. While the process may not be able to identify the exact section of the technical standard referenced by the IP service provider, it can point to a similar direction by identifying related parts, such as to a upper section of a referenced section, and the correct standard-essential claims. In addition, ***P5*** presents several opportunities arising from the use of semantic analysis for the assessment of standard-essentiality, such as the identification of potential and undeclared SEPs that fall under the definition of watchful waiting patents as defined by Herzberg et al. (2024).

The results of the logistic regression confirm the assumption that truly essential patents have a closer similarity to a technical standard than non-essential patents by showing that a higher DSS-Jaccard similarity is associated with a higher probability that patents are truly standard-essential. In addition, the number of similarities between a patent and the sections of a standard is related to the likelihood that the patent is truly standard-essential. This confirms recent literature which shows that SEPs tend to comprise more claims (Berger et al., 2012), as a higher number of claims is likely to produce to more similarities with a standard. Also, the process identifies many patents that appear to be standard-essential due to their similarity to the selected technical standard, but have not been declared as such at the ETSI. This result corroborates the existence of the watchful waiting strategy for SEPs as presented by Herzberg et al. (2024), which assumes that the declaration of standard-essentiality is intentionally or unintentionally delayed by the patent holder in order to eventually take competitors by surprise and thus gain a competitive advantage.

3.5.4 Implications for this dissertation

By presenting a new method that uses digital technologies, in this case, text-mining, to determine economic, legal, and technological patent quality, publication ***P5*** contributes to answering ***RQ-3*** of this dissertation, improving the holistic assessment of patent quality in patent intelligence for stakeholders such as IP & technology analysts.

As pointed out in ***P5***, it is crucial to determine the standard-essentiality of patents when relying on a technical standard. Companies that utilize the technology of a standard are required to license the related patents. However, since patents are not examined for standard-essentiality at the time of declaration and as there is a high rate of over- and under-declaration, the standard-essentiality of a patent must be verified by the company that is relying on the technical standard, which is a time-consuming task that requires legal and technological expertise. With the help of text-mining techniques, this workload can be reduced, patent claims can be compared both quantitatively and qualitatively with the sections of a technical standard, and an assessment of patentability can be made. However, several challenges arise for patent intelligence using text-mining for assessing the economic-legal-technological quality indicator of standard-essentiality.

The main challenge that results from the heterogeneity of the texts to be compared (La Justicia de Torre et al., 2018). Although this challenge is addressed by ***P5*** through harmonization of the language used in patents and technical standards, there may still be still differences that need to be considered in the future. Furthermore, traditional text-mining is fraught with the problem of ambiguity, as this type of technique is unable to recognize homonyms and synonyms of words (Arts et al., 2018; Gaikwad, Sonali, Vijay et al., 2014; Helmers et al., 2019). However, patents and technical standards often apply specific technical terminology to facilitate the use of a particular technology. This terminology is generally standardized and used uniformly throughout the relevant technical language. Consequently, this issue should be less pronounced in the analysis of patents and standards. Furthermore, in the context of technical standards, domain knowledge must be integrated (Akilan, 2015) in order to sift and analyze the results of the presented semantic process. Although the semantic process provides an assessment or rather a kind of probability, it ultimately takes a person skilled in the art to decide whether a patent is standard-essential or not and thereby validate the process.

Moreover, although the structure of technical standards is specified by guidelines (ETSI, 2020, 2022), technical standards may comprise numerous date formats such as tables, figures, formulas, etc. that cannot be extracted using conventional text-mining techniques. This can lead to false-negative assessments of patents, i.e. patents do not appear to be essential because they include content from this unstructured data (La Justicia de Torre et al., 2018).

4 Discussion

This section provides answers to the overarching research questions ***RQ-1**, **RQ-2**,* and ***RQ-3*** of this dissertation. ***RQ-1*** (*How can stakeholder theory be used to obtain a consistent understanding of patent quality?*) is answered by discussing the results of publication ***P1****.* To answer ***RQ-2*** (*How can an assessment of patent quality be conducted under consideration of the stakeholders involved in patent intelligence?*)*,* a mathematical expression of patent quality from the literature is applied to the identified stakeholders of patent intelligence. In answer to ***RQ-3*** (*How can digital technologies improve the assessment of patent quality for patent intelligence stakeholders and what challenges arise from their use?*), the challenges associated with digital technologies are summarized in a table and discussed.

4.1 Exploring the dimensions of patent quality

By proposing the consideration of different stakeholders in order to achieve a consistent understanding of patent quality, stakeholder theory shows that the concept of patent quality is difficult to define universally and must be interpreted depending on the perspective adopted. In the absence of a conclusive definition of patent quality, recent publications have attempted to align and/or define the terms of quality and value. Ananthraman et al. (2023), for example, combine the concepts of quality and value and categorize indicators for measuring patent quality and value into dimensions derived from the *ex ante* theory (for this and the following, see Perel, 2014). This theory analyses patent quality on the basis of four dimensions, which primarily take a legal perspective: Subject matter eligibility, utility, novelty and non-obviousness, and finally clarity and definiteness. Grimaldi and Cricelli (2020) show that indicators of patent value can be divided into the areas of legal, technology, market conditions, finance, and strategy. All of the above approaches have a common shortcoming: they do not regard both terms from different perspectives (Squicciarinii et al., 2023; Wu et al., 2021).

The consideration of different stakeholders and their perspectives is not entirely unusual: Guerrini (2014) identifies different stakeholders of patent quality and draws on them to describe the patent quality dimensions of probable validity, clarity, faithfulness, social utility, and commercial success. However, the article by Guerrini (2014) has two shortcomings: first, the dimensions are relatively specific and describe indicators rather than perspectives to be adopted, which makes it difficult to use them universally, e.g. to provide an overview of the indicators for these dimensions as done in publication ***P1***.[11] Second, the dimensions are not suitable for evaluating the quality of every patent. For instance, evaluating patents in emerging technologies such as quantum computing can be particularly challenging using Guerrini's dimensions, e.g. social utility is hard to gauge when the technology's impact on society is speculative and the societal benefits may not be immediately clear or measurable.

By applying stakeholder theory at the most fundamental level at which a patent can be considered, i.e. that of the patent as a *quid pro quo* (Walter and Schnittker, 2016), ***P1*** overcomes the aforementioned shortcomings and contributes to answering ***RQ-1*** by defining three basic perspectives on patent quality: Economic quality describes the ability of a patent to create value for the patent holder, influence markets and open up business opportunities. This quality is determined by external factors such as market demand and the competitive environment as well as internal factors such as the investment in the patent. Legal quality describes the ability of a patent to protect innovations and be legally enforceable. It includes aspects such as patentability, validity, and infringement potential. Technological quality describes the technological features of a patent and its social impact. This includes the originality and innovative character of a technical solution as well as its contribution to general technological development.

It should be noted, however, that the quality of a patent – at least considering the above definitions – cannot be equated with the quality of the invention it protects, as Guerrini (2014) mentions. Understanding this distinction is essential for a consistent understanding of patent quality, as a high-quality invention does not automatically result in a high-quality patent grant. The economic quality of an invention, for example, can be determined by the amount invested in its development (Chen et al., 2021). A patent grant has an economic quality if it creates commercial value. This can be achieved through licensing agreements (Erutku and Richelle, 2007; Lee, 2009) or the creation of new business opportunities (Wang and Hsieh, 2015). However, it is possible that the invention claimed in a patent is of high economic quality, e.g. if large investments have been made for its development, while the patent itself lacks economic quality, e.g. if it is not licensed and does not generate any profit.

[11] *Note:* For example, validity is a frequently used indicator for patent quality (e.g. Mann and Underweiser, 2012), and other frequently used indicators, such as scientific linkage (Trappey et al., 2012), are difficult to assign to the presented dimensions of (Guerrini, 2014).

Furthermore, the perception and evaluation of patent quality and the interpretation of its economic, legal, and technological dimensions depend on the hierarchical level at which a stakeholder is located in an organization. This means that different stakeholders at different hierarchical levels of a company or organization bring different criteria and perspectives to the evaluation of patent quality. For example, high-level decision-makers, such as the strategic management of a company, who are primarily concerned with comparing their patent portfolio to that of competitors, will have a different view of patent quality than stakeholders who are directly involved in the development of the invention and primarily focus on a single patent, such as the inventor.

In summary, patent quality based on ISO 9000 can be defined as *the degree to which a patent and/or patent portfolio – possibly irrespective of the claimed invention – fulfills the economic, legal, and/or technological requirements of various stakeholders, either independently or in comparison to one or more other patents.*

4.2 Assessing patent quality for patent intelligence stakeholders

To answer ***RQ-2***, the generally applicable patent quality formula by Guerrini (2014) is applied to the patent quality dimensions and stakeholders presented in this dissertation. The objective patent quality PQ can be simplified on a scale from low to high as follows:

$$PQ = PQ_E + PQ_L + PQ_T \, , \qquad \text{Eq. 1}$$

where PQ_E, PQ_L, and PQ_T are the economic, legal and technological patent quality. However, the quality of a patent is subjective. For example, a patent examiner verifies whether a patent application fulfills all legal requirements and ensures that the patent is drafted so that all information necessary to promote technological progress is disclosed. According to the definitions in ***P1***, the patent examiner is therefore interested in the legal and technological quality, while the economic quality of the patent to be examined is relatively unimportant. As suggested by Guerrini (2014), the weighting of the dimensions enables an individual adaption of Eq. 1 to a stakeholder so that a patent can be considered to be of high quality even if it only fulfills the requirements of single dimension. Accordingly, patent quality for a stakeholder of patent intelligence PQ_{ST} can be defined as:

$$PQ_{ST} = RI_{PQ_E} \cdot PQ_E + RI_{PQ_L} \cdot PQ_L + RI_{PQ_T} \cdot PQ_T, \qquad \text{Eq. 2}$$

where PQ_E, PQ_L, and PQ_T are the economic, legal and technological patent quality multiplied by their relative importance RI_{PQ_E}, RI_{PQ_L}, and RI_{PQ_T}. For patent examiners, who are not concerned with the economic quality of a patent, this would mean that the product of RI_{PQ_E} and PQ_E is zero. Using the indicators assigned to the dimensions of patent quality in publication ***P1***, the quality of one or more patents can be assessed using the above Eq. 2. This allows patent intelligence stakeholders to carry out a holistic benchmarking of patents. For example, strategic management could use the stock-market reaction to grant events (Kogan et al., 2017) (economic quality), the number of litigations (Crespi et al., 2007) (legal quality) and the number of forward citations (Trajtenberg, 1990) (technological quality) to assess the quality of its patent portfolio –adjusting the weighting of dimensions individually – and benchmark it against a competitor's patent portfolio to evaluate performance, at least if both are exchange-listed. For a more nuanced and complex benchmarking, indicators can be used for the intersections of the dimensions, such as generality for economic-technological quality.

4.3 Assessing patent quality using digital technologies

RQ-3 addresses the extent to which digital technologies can improve the assessment of patent quality in patent intelligence and what challenges are associated with this. As regards the use of digital technologies, an answer is given in each publication belonging to this dissertation. ***P1*** analyzes the current state-of-the-art and formulates research propositions, which are then addressed in publications ***P2***, ***P3***, ***P4,*** and ***P5***. In these publications, new methods are introduced to assess at least the legal quality of a patent, the quality that is represented in all publications. This is done in publication ***P2*** with a text-mining-based method to assess patentability, in publication ***P3*** with a text-mining-based method using machine translations to identify patent conflicts, in publication ***P4*** with a machine learning based method to assess statutory novelty, and in publication ***P5*** with a text-mining-based method to assess standard-essentiality.

An answer regarding the challenges arising from the use of digital technologies is provided by summarizing the main challenges associated with the application of text-mining and machine learning for patent quality assessment, as shown in **Table 3**. These challenges are derived from the implications of the individual publications for this dissertation. For example, the challenge of *data size* refers to the amount and scalability of data required for training accurate and robust machine learning models, as is the case in ***P4***.

Table 3: Challenges of using digital technologies for patent quality assessment in response to ***RQ-3***.

Challenge	***Brief description***	***Publications concerned***
Ambiguity	Ambiguity refers to the challenge of interpreting and analyzing text due to the multiple meanings of words and phrases (e.g. the bank of the river or the financial institution). This issue manifests in lexical (e.g., homonyms and polysemy), syntactic, and semantic forms. Ambiguity reduces accuracy, complicates processing, and can lead to noise in extracted data.	***P2, P3, P5***
Complexity of the configuration space	The challenge of configuration space complexity in deep learning models refers to the vast number of possible hyperparameter combinations and network architectures that make efficient model optimization and adaptation difficult.	***P4***
Data quality	Data quality refers to the challenge of accurately interpreting and analyzing text due to errors, inconsistencies, and incomplete information. Poor data quality, including misspellings, grammatical errors, and missing information, reduces accuracy and complicates processing.	***P2, P3, P4, P5***
Data size	The challenge of data size refers to the amount and scalability of data required for training accurate and robust models.	***P4***
Heterogeneity	Heterogeneity refers to the diversity and variability of textual data in terms of structure, style, language, and format. This complexity is caused by differences in grammar, vocabulary, writing conventions, and document types (e.g., tweets, academic papers, emails).	***P2, P3, P5***
Mistrust	Mistrust in machine learning is a significant challenge, referring to the lack of trust and confidence in the decisions and predictions made by machine learning models. This issue is caused by the non-transparency of complex algorithms or possible biases in training data.	***P4***
Need of domain knowledge	Domain knowledge is needed to obtain reliable results. Without expertise, it becomes difficult to accurately interpret context, disambiguate terms, and identify relevant patterns, resulting in less reliable and meaningful insights.	***P2, P3, P5***

Source: Author. *Note:* The challenges are listed in alphabetical order. For source information, refer to the individual implication sections.

Even though text-mining approaches have some disadvantages in comparison to machine learning, such as the problem of ambiguity, text-mining offers a significant advantage over machine learning approaches when it comes to evaluating the legal quality of a patent in patent intelligence: When legal aspects are evaluated in the context of legal quality assessment, it is crucial that stakeholders have confidence in the process and the resulting decisions. Despite its extensive capabilities, machine learning often encounters problems related to mistrust and tends to be perceived as a *black box* (Ribeiro et al., 2016), which affects the intention to use and accept the technology (Marangunić and Granić, 2015; Wu et al., 2011).[12] Li et al. (2008) show that users who are unfamiliar with or lack direct knowledge of a technology like machine learning, which may be the case among stakeholders such as patent attorneys or strategic management, must rely on the reputation and prestige of the technology or on those who support them when deciding whether to use it (McKnight et al., 1998).

Text-mining approaches, such as those used in publication ***P2*** and ***P5***, are particularly suitable for assessing legal quality. Semantic comparison, a text-mining method that is primarily used in the publications of this dissertation, provides a transparent and understandable method for the stakeholders involved in patent intelligence, which makes trust in and acceptance of the technology more likely. For example, if a patent is classified as statutorily novel through a semantic analysis procedure, e.g. by the process presented in ***P2***, the rationale for this decision is much more accessible and comprehensible to stakeholders than if a machine learning algorithm would make the same decision, e.g. by the deep learning model presented in ***P4***. This transparency is crucial for promoting trust in the assessment process. In addition, a text-mining approach does not require large data sets to implement, so it can be deployed on a smaller scale. This is a significant advantage over machine learning, which often requires large amounts of data and is typically more time and resource-intensive to implement, making text-mining more suitable for a *software-as-a-service* solution (Benlian and Hess, 2011; Schmitt and Denter, 2024).

In summary, it can be said that the implementation of text-mining provides a more transparent, understandable, and accessible method of assessing legal quality. Furthermore, text-mining can be effectively implemented without the extensive data requirements and resource investment associated with machine learning, making text-mining a more viable and confidence-inspiring solution for stakeholders involved in legal patent quality assessment. Nevertheless, machine learning does have its uses, for example, to obtain a quick initial assessment of a legal aspect on a large scale, such as the statutory novelty in publication ***P4***.

[12] *Note:* Mistrust can significantly impact key factors that determine the intention to adopt a technology, such as perceived risk, trust, and performance expectancy. This connection is discussed in various theoretical frameworks on technology acceptance, including the Technology Acceptance Model and its extenions, e.g. TAM2, or the Unified Theory of Acceptance and Use of Technology (Marangunić and Granić, 2015; Venkatesh, 2000; Venkatesh et al., 2012).

5 Conclusion

This section concludes the dissertation in three ways: First, a summary of the answers to the three research questions is provided. Second, limitations and suggestions for future research are presented. Third, implications for scholarship and management are outlined.

5.1 Summary of answers to the research questions

The assessment of patent quality represents an analysis that is carried out consciously or unconsciously by stakeholders of patent management and, correspondingly, patent intelligence. However, two general challenges arise in this context: the ambiguity of the definition of quality and the handling of a vast amount of patent data. In an attempt to address these challenges, three research questions are elaborated in this dissertation.

The first research question **RQ-1** (*How can stakeholder theory be used to obtain a consistent understanding of patent quality?*) is answered by this dissertation as follows:

- Stakeholder theory deepens the understanding of patent quality by taking into account the differentiated perspectives of the various stakeholders and shows how quality is created for an individual stakeholder.
- On an abstract level, three basic perspectives on patent quality can be derived from stakeholder theory: economic, legal, and technological quality. The economic quality of a patent assesses its ability to create monetary value for the patent holder, influence markets, and open up business opportunities, depending on external factors such as market demand and competition, as well as internal factors such as investment in the patent. The legal quality of a patent assesses its ability to protect innovations and ensure effective legal enforceability, taking into account patentability, validity, and infringement potential. The technological quality of a patent assesses its technological characteristics and its societal impact, including the originality and innovation of the technical solution and its contribution to general technological development.
- Patent quality can be defined as *the degree to which a patent and/or patent portfolio – possibly irrespective of the claimed invention – fulfills the economic, legal, and/or technological requirements of various stakeholders, either independently or in comparison to one or more other patents.*

The answers to the second research question ***RQ-2*** (*How can an assessment of patent quality be conducted under consideration of the stakeholders involved in patent intelligence?*) are:

- In their key tasks, patent intelligence stakeholders are required to adopt an economic, legal, and technological perspective in order to assess the quality of a patent.
- By identifying indicators for each quality dimension and intersections thereof, patent intelligence stakeholders are provided with indicators, e.g. patentability or novelty, for assessing the dimensions of patent quality either individually or in combination.
- By obtaining a formulaic expression of patent quality from an individual stakeholder's perspective, patent intelligence stakeholders can compare patents and perform multidimensional benchmarking.

Finally, the third research question ***RQ-3*** (*How can digital technologies improve the assessment of patent quality for patent intelligence stakeholders and what challenges arise from their use?*), is answered as follows:

- By analyzing the state-of-the-art in patent quality assessment and deriving a research agenda, possible ways to improve the assessment of patent quality are identified.
- By facilitating manual analysis, digital technologies such as text-mining or machine learning can improve the assessment of legal patent quality, a task involving indicators that usually require the expertise of a patent attorney or examiner.
- Even if digital technologies are not suitable for making absolute statements about the legal quality of a patent, they can at least provide an initial assessment for the indicators of patentability, patent conflict probability, statutory novelty, and standard-essentiality, and reduce the manual workload.
- The use of digital technologies poses several challenges for the assessment of patent quality, namely ambiguity, data size, complexity of the configuration space, data quality, data size, heterogeneity, mistrust, and the need for domain knowledge.
- Text-mining provides stakeholders with a transparent and understandable method for patent quality assessment, fostering trust and acceptance.

Consequently, stakeholders such as data providers, IP & technology analysts or patent attorneys can more effectively integrate and utilize these technologies in their workflows.

- Machine learning approaches, although capable of processing large data sets to assess patent quality, often face challenges related to transparency and trust, as they are perceived as “black boxes”. This mistrust can hinder their acceptance and adoption among stakeholders such as patent attorneys and strategic management, who need clear and understandable justifications for decisions. However, machine learning can still provide valuable initial assessments on a large scale, offering quick insights that can complement more detailed analyses.

5.2 Limitations and future research

Apart from the limitations arising from the individual publications, some others need to be mentioned in connection with this dissertation which offers potential for future research.

Firstly, the dissertation is based on a purely theoretical derivation. No empirical analyses such as surveys or qualitative interviews were conducted to investigate how patents are analyzed in practice by different stakeholders. Conducting surveys and interviews with stakeholders involved in patent intelligence, e.g. IP & technology analysts, could provide valuable insights to further improve the presented assessment framework. The same applies to the discussion of text-mining being more suitable for legal analysis than machine learning. Here too, surveys and interviews with stakeholders involved in patent intelligence would be appropriate in order to investigate this further and gain insights into which analyses can be carried out with machine learning and at what point trust is no longer sufficient. In this context, it would be interesting to investigate whether the challenges of mistrust are tied geographically. A comparison of the risk aversion of countries such as Germany and the United States based on the Hofstede cultural dimension of uncertainty avoidance, for example, shows that Germany is significantly less risk-affine, and therefore the challenges of mistrust might be more pronounced in Germany than in the US (Hofstede, 2011).

Secondly, the assessment of the legal patent quality of a patent is the objective and thematic focus of publications ***P2-P5***. The assessment of economic and technological patent quality received less attention and was only partially addressed by the assessment of standard-essentiality in publication ***P5***. Further research could improve the assessment of economic and technological patent quality in patent intelligence using digital technologies, identify the challenges involved, and contribute to the existing literature (e.g. Squicciarinii et al., 2023)

Thirdly, this dissertation does not address downstream tasks following the assessment of patent quality. These include explaining the underlying text-mining techniques or machine learning algorithms to potential stakeholders, communicating the derived findings, and involving the relevant stakeholders in the analysis process (Denter, 2022). Especially in the case of non-transparent models, subject to mistrust, comprehensive communication is necessary (Thiebes et al., 2021) to increase trust in the results and implement recommendations based on them (Mahmud et al., 2022).

Lastly, this dissertation does not examine whether patent quality assessment using text-mining or machine learning should be performed in-house or outsourced, nor to what extent such tasks should or should not be performed. The nature, industry, and size of an organization may determine the appropriate scope of this kind of assessment and the choice of digital technology (Denter, 2022). It is important to note that implementing of machine learning for patent quality assessment can be more challenging and may require a higher level of patent intelligence maturity (Moehrle et al., 2017), which makes machine learning more suitable for *software-as-a-service* (Benlian and Hess, 2011). Future research could benefit from embedding and outlining the individual methods in a maturity model according to their level of maturity as suggested by Denter (2022), addressing the requirements and aforementioned challenges associated with the use of digital technologies for assessing patent quality. It should account for the different maturity levels of machine learning technologies and the complexity of their implementation to guide organizations on which technologies and methods are suitable for their specific needs and maturity.

5.3 Implications for management and scholarship

In addition to the implications arising from each of the publications, this dissertation makes contribution to both management and scholarship by answering the research questions posed in several ways.

First, the dissertation shows how stakeholder theory can be applied to the assessment of patent quality. In doing so, the dissertation extends the theoretical scope of stakeholder theory by demonstrating its applicability not only to classic business problems (Parmar et al., 2010), but also to specific technological and legal topics such as patent quality. This extension contributes to a deeper theoretical understanding of stakeholder theory for both scholars and managers and opens up new areas of research for scholars in which this theory can be applied.

Second, by this extension, the dissertation advances the understanding of patent quality through the lens of stakeholder theory, thereby contributing to the substantial stream of literature the assessment and definition of patent quality (Ananthraman et al., 2023; e.g. Guerrini, 2014; Higham et al., 2021; Squicciarinii et al., 2023). By defining three basic perspectives of patent quality, a multidimensional concept of patent quality is introduced that captures the concept of patent value and quality and aligns the indicators for assessing the value and quality of a patent. By showing how different stakeholder interests and perspectives can be integrated, the dissertation also contributes to the literature by developing a more comprehensive and differentiated view of patent quality (Guerrini, 2014). This helps scholars to better capture the complexity and multidimensionality of patents and provides managers a basis for a more nuanced patent analysis.

Third, the dissertation contributes to the existing literature on the use of digital technologies in patent quality assessment emphasizing the necessity and relevance of text-mining approaches in the age of discriminative and generative artificial intelligence (Erdogan et al., 2024; e.g. Wu et al., 2016; Wu et al., 2021). In particular, it sheds light on the application of text-mining and machine learning in the context of legal patent quality and provides a starting point for future research. Machine learning models are subject to the problem of mistrust (Leung et al., 2016; Nunes and Jannach, 2017), which can be particularly pronounced in legal patent analysis, where wrong decisions may have serious consequences. Legal patent analysis must be meticulous and reliable because it decides whether an invention is patentable or whether it infringes an existing patent. Failure in this process, e.g. overlooking an existing patent that affects the novelty of an invention, poses a significant risk that affects the acceptance of the technology by the stakeholders involved (Li et al., 2008; Marangunić and Granić, 2015) and can lead to costly litigation, financial losses, and significant reputational damage (Lanjouw and Schankerman, 2004b). Machine learning models used for this type of analysis are often complex and lack transparency, which makes it difficult for stakeholders to rely on the results as they cannot fully understand how the decisions were made and how reliable they are. In contrast, text-mining offers a practicable approach, especially for small and medium-sized enterprises that do not have the same financial resources as large companies (Holgersson, 2013; Spithoven et al., 2013), as it can support legal patent analyses without the need to develop and train costly and complex machine learning models. Organizations wishing to employ machine learning for the assessment of legal patent quality could initially use text-mining and machine learning simultaneously to gain stakeholder trust in the technology and provide positive *second-hand* information as suggested by Li et al. (2008). Accordingly, this dissertation offers valuable insights and impulses for the further development and implementation of digital technologies in the field of patent quality assessment for scholars and managers, in particular regarding the legal quality of a patent, and beyond.

Fourth, particularly for managers, patent quality assessment can be made more efficient and accurate by applying computer-based methods using text-mining or machine learning. Manual reviews and assessments of patents are time-consuming and prone to bias and noise (Kahneman et al., 2021; Kovács, 2017). Computer-based procedures are able to perform these tasks faster and more consistently and can be used systematically for initial assessments. For example, the computer-based method presented in publication ***P2*** enables an assessment of patentability through mathematical-logical comparisons of patents, while the method presented in publication ***P4*** provides IP analysts with an initial assessment of the statutory novelty of a patent application. Such computer-based processes minimize the human bias that can occur in manual processes, such as those found with patent examiners (Kovács, 2017; Whalen, 2018). Furthermore, they offer scalability, so that large numbers of patents can be analyzed and compared, which is particularly beneficial for patent offices (e.g. for patent examiners analyzing the entire prior art) and companies (e.g. for strategic management analyzing complete patent portfolios).

Fifth, improved decision-making is enabled for managers and other parties involved in patent management, and thus patent intelligence. By considering a large numbers of patents to assess the quality of a patent portfolio, better and more informed decisions can be made, leading to improved patent strategy and management. Companies have the possibility to manage their patent portfolios in a more targeted way by better assessing which patents are of strategic value, e.g. because they are standard-essential. At the same time, a more nuanced assessment of patent quality can lead to an optimization of R&D investments, allowing strategic and product management to allocate resources to promising, patentable inventions. A more differentiated assessment of patent quality enables better informed licensing decisions, e.g. the specification of appropriate royalty rates based on legal quality, including factors such as standard-essentiality or validity.

Sixth, the dissertation contributes to the 7D patent management maturity model of Moehrle et al. (2017) by showing that the element of evaluation and valuation of the patent intelligence dimension does not only involve assessing technological and monetary quality but must also include an assessment of legal quality, which is crucial for managers. Furthermore, it appears that the quality of a patent or portfolio is also assessed in the elements of acquisition, business segment analysis, and prior art analysis. For example, patents of high technological quality can be identified for the purpose of gaining knowledge from them in the acquisition element, or the legal quality of a patent can be determined by examining its patentability in the element of prior art analysis.

References

Abbas, A., Zhang, L., Khan, S.U., 2014. A literature review on the state-of-the-art in patent analysis. World Patent Information 37, 3–13. doi:10.1016/j.wpi.2013.12.006.

Adams, R., Bessant, J., Phelps, R., 2006. Innovation management measurement: A review. International Journal of Management Reviews 8 (1), 21–47. doi:10.1111/j.1468-2370.2006.00119.x.

Agostini, L., Nosella, A., Teshome, M.B., 2019. Towards the development of scales to measure patent management. World Patent Information 58, 101909. doi:10.1016/j.wpi.2019.101909.

Akilan, A., 2015. Text mining: Challenges and future directions, in: 2015 2nd International Conference on Electronics and Communication Systems (ICECS). 2015 2nd International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India. 26.02.2015 - 27.02.2015. IEEE, pp. 1679–1684.

Albert, M.B., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Research Policy 20 (3), 251–259. doi:10.1016/0048-7333(91)90055-U.

Alderucci, D., Ashley, K., 2020. Using AI to Analyze Patent Claim Indefiniteness. IP Theory 9 (1).

Ananthraman, S., Cambré, B., Kittler, M., Delcamp, H., 2023. Divide and conquer: Relating patent quality and value in a conceptual framework based on a systematic review. International Journal of Management Reviews. doi:10.1111/ijmr.12354.

Aristodemou, L., 2020. Identifying Valuable Patents: A Deep Learning Approach. Apollo - University of Cambridge Repository.

Aristodemou, L., Tietze, F., 2018. The state-of-the-art on Intellectual Property Analytics (IPA): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. World Patent Information 55, 37–51. doi:10.1016/j.wpi.2018.07.002.

Arts, S., Cassiman, B., Gomez, J.C., 2018. Text matching to measure patent similarity. Strategic Management Journal 39 (1), 62–84. doi:10.1002/smj.2699.

Arts, S., Fleming, L., 2018. Paradise of Novelty—Or Loss of Human Capital? Exploring New Fields and Inventive Output. Organization Science 29 (6), 1074–1092. doi:10.1287/orsc.2018.1216.

Arts, S., Hou, J., Gomez, J.C., 2021. Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. Research Policy 50 (2), 104144. doi:10.1016/j.respol.2020.104144.

Asche, G., 2017. “80% of technical information found only in patents” – Is there proof of this [1] ? World Patent Information 48, 16–28. doi:10.1016/j.wpi.2016.11.004.

Ashtor, J.H., 2022. Modeling patent clarity. Research Policy 51 (2), 1–15. doi:10.1016/j.respol.2021.104415.

Aumayr, K.J., 2019. Erfolgreiches Produktmanagement. Springer Fachmedien Wiesbaden, Wiesbaden.

Bader, M.A., Gassmann, O., Ziegler, N., Ruether, F., 2012. Getting the most out of your IP--patent management along its life cycle. Drug discovery today 17 (7-8), 281–284. doi:10.1016/j.drudis.2011.10.025.

Baier, L., Jöhren, F., Seebacher. Stefan (Eds), 2019. CHALLENGES IN THE DEPLOYMENT AND OPERATION OF MACHINE LEARNING IN PRACTICE.

Baldini, N., Grimaldi, R., Sobrero, M., 2007. To patent or not to patent? A survey of Italian inventors on motivations, incentives, and obstacles to university patenting. Scientometrics 70 (2), 333–354. doi:10.1007/s11192-007-0206-5.

Banh, L., Strobel, G., 2023. Generative artificial intelligence. Electronic Markets 33 (1). doi:10.1007/s12525-023-00680-1.

Benlian, A., Hess, T., 2011. Opportunities and risks of software-as-a-service: Findings from a survey of IT executives. Decision Support Systems 52 (1), 232–246. doi:10.1016/j.dss.2011.07.007.

Berger, F., Blind, K., Thumm, N., 2012. Filing behaviour regarding essential patents in industry standards. Research Policy 41 (1), 216–225. doi:10.1016/j.respol.2011.07.004.

Blei, D.M., Lafferty, J.D., 2006. Dynamic topic models, in: Proceedings of the 23rd international conference on Machine learning - ICML '06. the 23rd international conference, Pittsburgh, Pennsylvania. 25.06.2006 - 29.06.2006. ACM Press, New York, New York, USA, pp. 113–120.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet Allocation. Journal of machine learning research (3), 993–1022.

Blenk, A., Kalmbach, P., Kellerer, W., Schmid, S., 2017. o'zapft is, in: Proceedings of the Workshop on Big Data Analytics and Machine Learning for Data Communication Networks. SIGCOMM '17: ACM SIGCOMM 2017 Conference, Los Angeles CA USA. 21 08 2017 21 08 2017. ACM, New York, NY, USA, pp. 19–24.

Block, C., Wustmans, M., Laibach, N., Bröring, S., 2021. Semantic bridging of patents and scientific publications – The case of an emerging sustainability-oriented technology. Technological Forecasting and Social Change 167, 120689. doi:10.1016/j.techfore.2021.120689.

Bloom, N., van Reenen, J., 2002. Patents, Real Options and Firm Performance. The Economic Journal 112 (478), C97-C116. doi:10.1111/1468-0297.00022.

Bosse, D.A., Phillips, R.A., Harrison, J.S., 2009. Stakeholders, reciprocity, and firm performance. Strategic Management Journal 30 (4), 447–456. doi:10.1002/smj.743.

Bourgeois, L.J., 1980. Strategy and Environment: A Conceptual Integration. Academy of Management Review 5 (1), 25. doi:10.2307/257802.

Boutaba, R., Salahuddin, M.A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., et al., 2018. A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. Journal of Internet Services and Applications 9 (1). doi:10.1186/s13174-018-0087-2.

Boutellier, R., Gassmann, O., Zedtwitz, M. von, 2008. Managing Global Innovation. Springer Berlin Heidelberg, Berlin, Heidelberg.

Brachtendorf, L., Gaessler, F., Harhoff, D., 2023. Truly standard-essential patents? A semantics-based analysis. Journal of Economics & Management Strategy 32 (1), 132–157. doi:10.1111/jems.12500.

Brodley, C.E., Rebbapragada, U., Small, K., Wallace, B.C., 2012. Challenges and Opportunities in Applied Machine Learning. AI Magazine 33 (1), 11–24. doi:10.1609/aimag.v33i1.2367.

Brynjolfsson, E., Li, D., Raymond, L., 2023. Generative AI at Work, Cambridge, MA.

Brynjolfsson, E., McAfee, A., 2014. The second machine age : work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company, New York.

Brynjolfsson, E., Mitchell, T., 2017. What can machine learning do? Workforce implications. Science (New York, N.Y.) 358 (6370), 1530–1534. doi:10.1126/science.aap8062.

Camarota, A., 2016. The Pillars of Patent Quality. Technology & Innovation 18 (1), 75–77. doi:10.21300/18.1.2016.75.

Cao, Y., Zhao, L., 2013. Analysis of patent management effects on technological innovation performance. Baltic Journal of Management 8 (3), 286–305. doi:10.1108/BJOM-May-2012-0033.

Cao, Z., Zhao, H., 2008. Research of Knowledge Acquisition and Modeling Method Based on Patent Map, in: 2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop. 2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop (KAM 2008 Workshop), Wuhan, China. 21.12.2008 - 22.12.2008. IEEE, pp. 1090–1094.

Castelvecchi, D., 2016. Can we open the black box of AI? Nature 538 (7623), 20–23. doi:10.1038/538020a.

Caviggioli, F., 2016. Technology fusion: Identification and analysis of the drivers of technology convergence using patent data. Technovation 55-56, 22–32. doi:10.1016/j.technovation.2016.04.003.

Chen, Y.-S., Chang, K.-C., 2009. Using neural network to analyze the influence of the patent performance upon the market value of the US pharmaceutical companies. Scientometrics 80 (3), 637–655. doi:10.1007/s11192-009-2095-2.

Chen, Z., Zhang, J., Zi, Y., 2021. A cost-benefit analysis of R&D and patents: Firm-level evidence from China. European Economic Review 133, 103633. doi:10.1016/j.euroecorev.2020.103633.

Chen, Z.-H., Zhou, Z.-J., 2018. Research on the Evaluation Index System for Patent Financing Capacity of High-tech Enterprises, in: Proceedings of the 2018 4th International Conference on Social Science and Higher Education (ICSSHE 2018). Proceedings of the 2018 4th International Conference on Social Science and Higher Education (ICSSHE 2018), Sanya, China. 28.09.2018 - 30.09.2018. Atlantis Press, Paris, France.

Chikkamath, R., Endres, M., Bayyapu, L., Hewel, C., 2020. An Empirical Study on Patent Novelty Detection: A Novel Approach Using Machine Learning and Natural Language Processing, in: 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS). 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS), Paris, France. 14.12.2020 - 16.12.2020. IEEE, pp. 1–7.

Chitale, S., Lawler, C., Macfarlane, S., 2020. Understanding the basics of patenting. Nature biotechnology 38 (3), 263–270. doi:10.1038/s41587-020-0447-x.

Cho, R.L.-T., Liu, J.S., Ho, M.H.-C., 2023. Exploring a Patent's Essentiality to the HEVC Standard: A Retrospective View. IEEE Transactions on Engineering Management 70 (12), 4035–4047. doi:10.1109/TEM.2021.3111963.

Chon, K.-S., Olsen, M.D., 1990. Applying the strategic management process in the management of tourism organizations. Tourism Management 11 (3), 206–213. doi:10.1016/0261-5177(90)90043-9.

Choudhury, P., Starr, E., Agarwal, R., 2020. Machine learning and human capital complementarities: Experimental evidence on bias mitigation. Strategic Management Journal 41 (8), 1381–1411. doi:10.1002/smj.3152.

Chung, J., Ko, N., Kim, H., Yoon, J., 2021. Inventor profile mining approach for prospective human resource scouting. Journal of Informetrics 15 (1), 101103. doi:10.1016/j.joi.2020.101103.

Clarke, N.S., 2018. The basics of patent searching. World Patent Information 54, 4-10. doi:10.1016/j.wpi.2017.02.006.

Colangelo, G., Aguggia, A., 2023. SEPs Infringement and Competition Law Defence in German Case Law. SSRN Electronic Journal. doi:10.2139/ssrn.4444068.

Conley, J.G., Bican, P.M., Ernst, H., 2013. Value Articulation: A Framework for the Strategic Management of Intellectual Property. California Management Review 55 (4), 102–120. doi:10.1525/cmr.2013.55.4.102.

Contreras, J.L., 2013. TECHNICAL STANDARDS AND 'EX ANTE' DISCLOSURE: RESULTS AND ANALYSIS OF AN EMPIRICAL STUDY. Jurimetrics (vol. 53, no. 2), 163–211.

Crespi, G.A., Geuna, A., Nesta, L., 2007. The mobility of university inventors in Europe. The Journal of Technology Transfer 32 (3), 195–215. doi:10.1007/s10961-006-9012-0.

Dang, J., Motohashi, K., 2015. Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality. China Economic Review 35, 137–155. doi:10.1016/j.chieco.2015.03.012.

David, F.R., 2011. Strategic management: Concepts and cases (1Third edition). Pretince Hall; Pearson, Upper Saddle River, New Jersey, Boston, Massachusetts.

DeCenzo, D.A., Robbins, S.P., 2005. Fundamentals of human resource management (8e éd. 2005). Wiley, New Jersey.

Denter, N., 2022. Machine learning for patent intelligence: opportunities and challenges.

Denter, N., Caferoglu, H., Moehrle, M.G., 2019. Applying Dynamic Topic Modeling for Understanding the Evolution of the RFID Technology, in: Technology management in the world of intelligent systems. Proceedings : PICMET '19 : Portland International Conference on Management of Engineering and Technology. 2019 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA. 8/25/2019 - 8/29/2019. IEEE, Piscataway, NJ, pp. 1–9.

Denter, N.M., Aaldering, L.J., Caferoglu, H., 2022. Forecasting future bigrams and promising patents: introducing text-based link prediction. foresight. doi:10.1108/FS-03-2021-0078.

Denter, N.M., Seeger, F., Moehrle, M.G., 2023. How can Blockchain technology support patent management? A systematic literature review. International Journal of Information Management 68, 102506. doi:10.1016/j.ijinfomgt.2022.102506.

Dessler, G., 2013. Human resource management (Thirteenth edition). Pearson Education Inc, Boston.

Dhulap, S., Joshi, A., Kulkarni, M.G., 2015. Obvious to try and non-obviousness post KSR: origins and implications. International Journal of Intellectual Property Management 8 (3/4), 190–206. doi:10.1504/IJIPM.2015.076548.

Dirnberger, D., 2011. A guide to efficient keyword, sequence and classification search strategies for biopharmaceutical drug-centric patent landscape searches - A human recombinant insulin patent landscape case study. World Patent Information 33 (2), 128–143. doi:10.1016/j.wpi.2010.12.003.

Donaldson, T., Preston, L.E., 1995. The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. Academy of Management Review 20 (1), 65. doi:10.2307/258887.

DPMA, 2022. Annual Report 2022 (downloaded on 11 April 2024 from https://www.dpma.de/docs/english/jahresberichte/dpma_jb2022_engl_nichtbarrpdf.pdf).

Du, M., Liu, N., Hu, X., 2019. Techniques for interpretable machine learning. Communications of the ACM 63 (1), 68–77. doi:10.1145/3359786.

Dybå, T., Dingsøyr, T., 2008. Strength of evidence in systematic reviews in software engineering, in: Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement. ESEM '08: 2008 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, Kaiserslautern Germany. 09 10 2008 10 10 2008. ACM, New York, NY, USA, pp. 178–187.

EPO, 2020a. The European Patent Convention: Article 54 Novelty (downloaded on 3 January 2022 from https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ar54.html).

EPO, 2020b. The European Patent Convention: Article 56 Inventive Step (downloaded on 3 January 2022 from https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ar56.html).

Erdogan, Z., Altuntas, S., Dereli, T., 2024. Predicting Patent Quality Based on Machine Learning Approach. IEEE Transactions on Engineering Management 71, 3144–3157. doi:10.1109/TEM.2022.3207376.

Ernst, H., 2003. Patent information for strategic technology management. World Patent Information 25 (3), 233–242. doi:10.1016/S0172-2190(03)00077-2.

Ernst, H., Conley, J., Omland, N., 2016. How to create commercial value from patents: the role of patent management. R&D Management 46 (S2), 677–690. doi:10.1111/radm.12210.

Erutku, C., Richelle, Y., 2007. Optimal Licensing Contracts and the Value of a Patent. Journal of Economics & Management Strategy 16 (2), 407–436. doi:10.1111/j.1530-9134.2007.00144.x.

Erzurumlu, S.S., Pachamanova, D., 2020. Topic modeling and technology forecasting for assessing the commercial viability of healthcare innovations. Technological Forecasting and Social Change 156, 120041. doi:10.1016/j.techfore.2020.120041.

ETSI, 2020. Writing World Class Standards, 1–33.

ETSI, 2022. Principles for drafting ETSI deliverables with the use of skeletons, 1–24.

Fan, S., Liu, G., Tu, Y., Zhu, J., Zhang, P., Tian, Z., 2023. Improved multi-criteria decision making method integrating machine learning for patent competitive potential Evaluation：A case study in water pollution abatement technology. Journal of Cleaner Production 403, 136896. doi:10.1016/j.jclepro.2023.136896.

Farrell, J., Shapiro, C., 2008. How Strong Are Weak Patents? American Economic Review 98 (4), 1347–1369. doi:10.1257/aer.98.4.1347.

Fattori, M., Pedrazzi, G., Turra, R., 2003. Text mining applied to patent mapping: a practical business case. World Patent Information 25 (4), 335–342. doi:10.1016/S0172-2190(03)00113-3.

Fawareh, H.M.A., Jusoh, S., Osman, W.R.S., 2008. Ambiguity in text mining, in: 2008 International Conference on Computer and Communication Engineering. 2008 International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, Malaysia. 13.05.2008 - 15.05.2008. IEEE, pp. 1172–1176.

Feldman, R., Sanger, J., 2009. The Text Mining Handbook. Cambridge University Press.

Fletcher, J. (Ed), 1992. Quality and risk assessment in patent searching and analysis.: Recent advances in chemical information II.

Foglia, P., 2007. Patentability search strategies and the reformed IPC: A patent office perspective. World Patent Information 29 (1), 33–53. doi:10.1016/j.wpi.2006.08.002.

Freeman, 1984. Strategic Management: A Stakeholder Approach. Cambridge University Press.

Freeman, R.E., 1994. The Politics of Stakeholder Theory: Some Future Directions. Business Ethics Quarterly 4 (4), 409–421. doi:10.2307/3857340.

Freeman, R.E., Harrison, J.S., Wicks, A.C., 2007. Managing for Stakeholders. Yale University Press.

Freunek, M., Bodmer, A., 2021. BERT based freedom to operate patent analysis.

Gaikwad, Sonali, Vijay, Chaugule, A., Patil, P., 2014. Text Mining Methods and Techniques. International Journal of Computer Applications 85 (17), 42–45. doi:10.5120/14937-3507.

Gassmann, O., Bader, M.A., Thompson, M.J., 2021a. Evaluating and Valuing Patents, in: Gassmann, O., Bader, M.A., Thompson, M.J. (Eds), Patent Management. Springer International Publishing, Cham, pp. 51–93.

Gassmann, O., Bader, M.A., Thompson, M.J. (Eds), 2021b. Patent Management. Springer International Publishing, Cham.

Gerken, J.M., Moehrle, M.G., 2012. A new instrument for technology monitoring: novelty in patents measured by semantic patent analysis. Scientometrics 91 (3), 645–670. doi:10.1007/s11192-012-0635-7.

Graham S., Hall B., Harhoff D., Mowery D, 2003. Patents in the Knowledge-Based Economy: Patent quality control: A comparison of U.S. patent re-examinations and European patent oppositions. National Academies Press, Washington, D.C.

Gregory, R.W., Henfridsson, O., Kaganer, E., Kyriakou, S.H., 2021. The Role of Artificial Intelligence and Data Network Effects for Creating User Value. Academy of Management Review 46 (3), 534–551. doi:10.5465/amr.2019.0178.

Gretzel, U., 2011. Intelligent systems in tourism. Annals of Tourism Research 38 (3), 757–779. doi:10.1016/j.annals.2011.04.014.

Griffiths, A., 2024. How to become an Intellectual Property or patent analyst (downloaded on 23 April 2024 from https://sourceadvisors.co.uk/insights/knowledge-hub/how-to-become-an-intellectual-property-analyst/).

Grimaldi, M., Cricelli, L., 2020. Indexes of patent value: a systematic literature review and classification. Knowledge Management Research & Practice 18 (2), 214–233. doi:10.1080/14778238.2019.1638737.

Guerrini, C.J., 2014. Defining Patent Quality. Fordham L. Rev. Vol. 82 (6).

Hackl, V., Müller, A.E., Granitzer, M., Sailer, M., 2023. Is GPT-4 a reliable rater? Evaluating consistency in GPT-4's text ratings. Frontiers in Education 8. doi:10.3389/feduc.2023.1272229.

Hall, B.H., Graham, S., Harhoff, D., Mowery, D.C., 2004. Prospects for Improving U.S. Patent Quality via Postgrant Opposition. Innovation Policy and the Economy (4), 115–143.

Hall, B.H., Jaffe, A., Trajtenberg, M., 2005. Market Value and Patent Citations. The RAND Journal of Economics 36 (1), 16–38.

Hall, B.H., MacGarvie, M., 2010. The private value of software patents. Research Policy 39 (7), 994–1009. doi:10.1016/j.respol.2010.04.007.

Hall, R., 1992. The strategic analysis of intangible resources. Strategic Management Journal 13 (2), 135–144. doi:10.1002/smj.4250130205.

Han, E.J., Sohn, S.Y., 2015. Patent valuation based on text mining and survival analysis. The Journal of Technology Transfer 40 (5), 821–839. doi:10.1007/s10961-014-9367-6.

Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation Frequency and the Value of Patented Inventions. Review of Economics and Statistics 81 (3), 511–515. doi:10.1162/003465399558265.

Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Research Policy 32 (8), 1343–1363. doi:10.1016/S0048-7333(02)00124-5.

Harrison, J.S., Wicks, A.C., 2013. Stakeholder Theory, Value, and Firm Performance. Business Ethics Quarterly 23 (1), 97–124. doi:10.5840/beq20132314.

Helmers, L., Horn, F., Biegler, F., Oppermann, T., Müller, K.-R., 2019. Automating the search for a patent's prior art with a full text similarity search. PloS one 14 (3), e0212103. doi:10.1371/journal.pone.0212103.

Herzberg, A., Denter, N.M., Moehrle, M.G., 2024. The watchful waiting strategy in standard-essential patents: The case of 5G technology. IEEE Transactions on Engineering Management, 1–50. doi:10.1109/TEM.2024.3374878.

Higham, K., Rassenfosse, G. de, Jaffe, A.B., 2021. Patent Quality: Towards a Systematic Framework for Analysis and Measurement. Research Policy 50 (4), 104215. doi:10.1016/j.respol.2021.104215.

Hofstede, G., 2011. Dimensionalizing Cultures: The Hofstede Model in Context. Online Readings in Psychology and Culture 2 (1). doi:10.9707/2307-0919.1014.

Holgersson, M., 2013. Patent management in entrepreneurial SMEs: a literature review and an empirical study of innovation appropriation, patent propensity, and motives. R&D Management 43 (1), 21–36. doi:10.1111/j.1467-9310.2012.00700.x.

Hotho, A., Nürnberger, A., Paaß, G., 2005. A Brief Survey of Text Mining. Journal for Language Technology and Computational Linguistics 20 (1), 19–62. doi:10.21248/jlcl.20.2005.68.

Imai, S., Lin, C.-W., Watada, J., Tzeng, G.-H., 2008. Knowledge acquisition in human resource management based on rough sets, in: PICMET '08 - 2008 Portland International Conference on Management of Engineering & Technology. Technology, Cape Town, South Africa. 27.07.2008 - 31.07.2008. IEEE, pp. 969–974.

International Trade Center, 2012. What is quality?

Janiesch, C., Zschech, P., Heinrich, K., 2021. Machine learning and deep learning. Electronic Markets 31 (3), 685–695. doi:10.1007/s12525-021-00475-2.

Jeon, D., Ahn, J.M., Kim, J., Lee, C., 2022. A doc2vec and local outlier factor approach to measuring the novelty of patents. Technological Forecasting and Social Change 174, 121294. doi:10.1016/j.techfore.2021.121294.

Jin, Y., Wang, C., 2022. Chinese court rules for the first time that it has jurisdiction over SEP global licensing disputes. 1747-1532 17 (2), 81–82. doi:10.1093/jiplp/jpab180.

Jo, T., 2019. Text Mining. Springer International Publishing, Cham.

Kahneman, D., Sibony, O., Sunstein, C.R., 2021. Noise: A flaw in human judgment (First edition). Little Brown Spark Hachette Book Group, New York, Boston, London.

Kaplan, S., Vakili, K., 2015. The double-edged sword of recombination in breakthrough innovation. Strategic Management Journal 36 (10), 1435–1457. doi:10.1002/smj.2294.

Kim, S., Yoon, B., 2021. Patent infringement analysis using a text mining technique based on SAO structure. Computers in Industry 125, 103379. doi:10.1016/j.compind.2020.103379.

KIPO, 2006. Application Procedure for Patents and Utility models: Substantial Examination (downloaded on 8 June 2022 from https://www.kipo.go.kr/en/HtmlApp?c=30101&catmenu=ek03_02_01).

Kitchenham, B., Brereton, P., 2013. A systematic review of systematic review process research in software engineering. Information and Software Technology 55 (12), 2049–2075. doi:10.1016/j.infsof.2013.07.010.

Knight, H.J., 2013. Patent strategy for researchers and research managers (Third ed.). Wiley, Chichester, West sussex, United Kingdom.

Kocheturov, A., Pardalos, P.M., Karakitsiou, A., 2019. Massive datasets and machine learning for computational biomedicine: trends and challenges. Annals of Operations Research 276 (1-2), 5–34. doi:10.1007/s10479-018-2891-2.

Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological Innovation, Resource Allocation, and Growth*. The Quarterly Journal of Economics 132 (2), 665–712. doi:10.1093/qje/qjw040.

Kovács, B., 2017. Too hot to reject: The effect of weather variations on the patent examination process at the United States Patent and Trademark Office. Research Policy 46 (10), 1824–1835. doi:10.1016/j.respol.2017.08.010.

Kowalski, T.J., Maschio, A., Megerditchian, S.H., 2003. Dominating global intellectual property: Overview of patentability in the USA, Europe and Japan. Journal of Commercial Biotechnology 9 (4). doi:10.5912/jcb41.

Krajec, R., 2020. Why Your Patent Attorney Does Not Want Your Patent To Be Granted. BlueIron IP https://blueironip.com/why-your-patent-attorney-does-not-want-your-patent-to-be-granted/).

Kühl, N., Hirt, R., Baier, L., Schmitz, B., Satzger, G., 2021. How to Conduct Rigorous Supervised Machine Learning in Information Systems Research: The Supervised Machine Learning Report Card. Communications of the Association for Information Systems 48 (1), 589–615. doi:10.17705/1CAIS.04845.

Kühl, N., Schemmer, M., Goutier, M., Satzger, G., 2022. Artificial intelligence and machine learning. Electronic Markets 32 (4), 2235–2244. doi:10.1007/s12525-022-00598-0.

La Justicia de Torre, C., Sánchez, D., Blanco, I., Martín-Bautista, M.J., 2018. Text Mining: Techniques, Applications, and Challenges. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 26 (04), 553–582. doi:10.1142/S0218488518500265.

Lambert, L.S., Newman, D.A., 2022. Construct Development and Validation in Three Practical Steps: Recommendations for Reviewers, Editors, and Authors*. Organizational Research Methods, 109442812211153. doi:10.1177/10944281221115374.

Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33 (1), 159. doi:10.2307/2529310.

Lanjouw, J.O., 1998. Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value. The Review of Economic Studies 65 (4), 671–710. doi:10.1111/1467-937X.00064.

Lanjouw, J.O., Schankerman, M., 2004a. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. The Economic Journal 114 (495), 441–465. doi:10.1111/j.1468-0297.2004.00216.x.

Lanjouw, J.O., Schankerman, M., 2004b. Protecting Intellectual Property Rights: Are Small Firms Handicapped? The Journal of Law and Economics 47 (1), 45–74. doi:10.1086/380476.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. doi:10.1038/nature14539.

Lee, C., Kang, B., Shin, J., 2015. Novelty-focused patent mapping for technology opportunity analysis. Technological Forecasting and Social Change 90, 355–365. doi:10.1016/j.techfore.2014.05.010.

Lee, C., Kwon, O., Kim, M., Kwon, D., 2018. Early identification of emerging technologies: A machine learning approach using multiple patent indicators. Technological Forecasting and Social Change 127, 291–303. doi:10.1016/j.techfore.2017.10.002.

Lee, J.-S., Hsiang, J., 2020. Patent claim generation by fine-tuning OpenAI GPT-2. World Patent Information 62, 101983. doi:10.1016/j.wpi.2020.101983.

Lee, M., Kim, S., Kim, H., Lee, J., 2022. Technology Opportunity Discovery using Deep Learning-based Text Mining and a Knowledge Graph. Technological Forecasting and Social Change 180, 121718. doi:10.1016/j.techfore.2022.121718.

Lee, S., Yoon, B., Park, Y., 2009. An approach to discovering new technology opportunities: Keyword-based patent map approach. Technovation 29 (6-7), 481–497. doi:10.1016/j.technovation.2008.10.006.

Lee, Y.-G., 2009. What affects a patent's value? An analysis of variables that affect technological, direct economic, and indirect economic value: An exploratory conceptual approach. Scientometrics 79 (3), 623–633. doi:10.1007/s11192-007-2020-5.

Lemley, M.A., Myhrvold, N., 2007. How to Make a Patent Market. Hofstra Law Review (Vol. 36: Iss. 2).

Lemley, M.A., Sampat, B., 2012. Examiner Characteristics and Patent Office Outcomes. Review of Economics and Statistics 94 (3), 817–827. doi:10.1162/REST_a_00194.

Leung, M.K.K., Delong, A., Alipanahi, B., Frey, B.J., 2016. Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. Proceedings of the IEEE 104 (1), 176–197. doi:10.1109/JPROC.2015.2494198.

Li, S., Zhang, X., Xu, H., Fang, S., Garces, E., Daim, T., 2020. Measuring strategic technological strength :Patent Portfolio Model. Technological Forecasting and Social Change 157, 120119. doi:10.1016/j.techfore.2020.120119.

Li, X., Hess, T.J., Valacich, J.S., 2008. Why do we trust new technology? A study of initial trust formation with organizational information systems. The Journal of Strategic Information Systems 17 (1), 39–71. doi:10.1016/j.jsis.2008.01.001.

Lippman, S.A., Rumelt, R.P., 2003. A bargaining perspective on resource advantage. Strategic Management Journal 24 (11), 1069–1086. doi:10.1002/smj.345.

List, J., 2012. Review of machine translation in patents – Implications for search. World Patent Information 34 (3), 193–195. doi:10.1016/j.wpi.2012.05.005.

Mahmud, H., Islam, A.N., Ahmed, S.I., Smolander, K., 2022. What influences algorithmic decision-making? A systematic literature review on algorithm aversion. Technological Forecasting and Social Change 175, 121390. doi:10.1016/j.techfore.2021.121390.

Malewicki, D., Sivakumar, K., 2004. Patents and product development strategies: a model of antecedents and consequences of patent value. European Journal of Innovation Management 7 (1), 5–22. doi:10.1108/14601060410515600.

Mann, R.J., Underweiser, M., 2012. A New Look at Patent Quality: Relating Patent Prosecution to Validity. Journal of Empirical Legal Studies 9 (1), 1–32. doi:10.1111/j.1740-1461.2011.01245.x.

Marangunić, N., Granić, A., 2015. Technology acceptance model: a literature review from 1986 to 2013. Universal Access in the Information Society 14 (1), 81–95. doi:10.1007/s10209-014-0348-1.

Marco, A.C., Sarnoff, J.D., deGrazia, C.A., 2019. Patent claims and patent scope. Research Policy 48 (9), 103790. doi:10.1016/j.respol.2019.04.014.

Marco, A.C., Toole, A.A., Miller, R., Frumkin, J., 2017. USPTO Patent Prosecution and Examiner Performance Appraisal. SSRN Electronic Journal. doi:10.2139/ssrn.2995674.

Marttin, E., Derrien, A.-C., 2018. How to apply examiner search strategies in Espacenet. A case study. World Patent Information 54, 33-43. doi:10.1016/j.wpi.2017.06.001.

McKnight, D.H., Cummings, L.L., Chervany, N.L., 1998. Initial Trust Formation in New Organizational Relationships. Academy of Management Review 23 (3), 473–490. doi:10.5465/AMR.1998.926622.

Miric, M., Jia, N., Huang, K.G.-L., 2022. Using Supervised Machine Learning to Create Categorical Variables for Use in Management Research: The Case for Identifying Artificial Intelligence Patents. Strategic Management Journal. doi:10.1002/smj.3441.

Moehrle, M.G., Caferoglu, H., 2019. Technological speciation as a source for emerging technologies. Using semantic patent analysis for the case of camera technology. Technological Forecasting and Social Change 146, 776–784. doi:10.1016/j.techfore.2018.07.049.

Moehrle, M.G., Gerken, J.M., 2012. Scientometrics 91 (3), 805–826. doi:10.1007/s11192-012-0682-0.

Moehrle, M.G., Walter, L., Geritz, A., Muller, S., 2005. Patent-based inventor profiles as a basis for human resource decisions in research and development. R&D Management 35 (5), 513–524. doi:10.1111/j.1467-9310.2005.00408.x.

Moehrle, M.G., Walter, L., Wustmans, M., 2017. Designing the 7D patent management maturity model – A capability based approach. World Patent Information 50, 27–33. doi:10.1016/j.wpi.2017.08.003.

Moehrle, M.G., Walter, L., Wustmans, M., 2018. Patente managen mit dem 7D Reifegradmodell: Erfassung - Bewertung - Verbesserung (1. Auflage). i3 - Management von Invention, Innovation, Information, [Bremen].

Moeller, A., Moehrle, M.G., 2015. Completing keyword patent search with semantic patent search: introducing a semiautomatic iterative method for patent near search based on semantic similarities. Scientometrics 102 (1), 77–96. doi:10.1007/s11192-014-1446-9.

Narin, F., 1994. Patent bibliometrics. Scientometrics 30 (1), 147–155. doi:10.1007/BF02017219.

Niemann, H., 2015. Corporate Foresight mittels Geschäftsprozesspatenten. Springer Fachmedien Wiesbaden, Wiesbaden.

Niemann, H., Moehrle, M.G., Frischkorn, J., 2017. Use of a new patent text-mining and visualization method for identifying patenting patterns over time: Concept, method and test application. Technological Forecasting and Social Change 115, 210–220. doi:10.1016/j.techfore.2016.10.004.

Nunes, I., Jannach, D., 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. User Modeling and User-Adapted Interaction 27 (3-5), 393–444. doi:10.1007/s11257-017-9195-0.

Oh, S., Choi, J., Ko, N., Yoon, J., 2020. Predicting product development directions for new product planning using patent classification-based link prediction. Scientometrics 125 (3), 1833–1876. doi:10.1007/s11192-020-03709-w.

Olavsrud, T., 2022. What is business analytics? Using data to improve business outcomes (downloaded on 23 April 2024 from https://www.cio.com/article/191157/what-is-business-analytics-using-data-to-predict-business-outcomes.html).

Özbilgin, M., Groutsis, D., Harvey, W.S. (Eds), 2014. International human resource management. Cambridge University Press, Melbourne, VIC, Australia, New York.

Ozdemir, S., Carlos Fernandez de Arroyabe, J., Sena, V., Gupta, S., 2023. Stakeholder diversity and collaborative innovation: Integrating the resource-based view with stakeholder theory. Journal of Business Research 164, 113955. doi:10.1016/j.jbusres.2023.113955.

Pakes, A., Schankerman, M., 1984. The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources. R&D, Patents and Productivity, 73–88.

Park, H., Ree, J.J., Kim, K., 2013. Identification of promising patents for technology transfers using TRIZ evolution trends. Expert Systems with Applications 40 (2), 736–743. doi:10.1016/j.eswa.2012.08.008.

Parmar, B.L., Freeman, R.E., Harrison, J.S., Wicks, A.C., Purnell, L., Colle, S. de, 2010. Stakeholder Theory: The State of the Art. The Academy of Management Annals 4 (1), 403–445. doi:10.1080/19416520.2010.495581.

Perel, M.F., 2014. An Ex Ante Method of Patent Valuation: Transforming Patent Quality into Patent Value. Journal of High Technology Law.

Pérez, J., Díaz, J., Garcia-Martin, J., Tabuenca, B., 2020. Systematic literature reviews in software engineering—enhancement of the study selection process using Cohen's

Kappa statistic. Journal of Systems and Software 168, 110657. doi:10.1016/j.jss.2020.110657.
Petralia, S., 2020. Mapping general purpose technologies with patent data. Research Policy 49 (7), 104013. doi:10.1016/j.respol.2020.104013.
Plečnik, J.M., Yang, L.L., Zhang, J.H., 2022. Corporate innovation and future earnings: does early patent disclosure matter? Accounting & Finance 62 (S1), 2011–2056. doi:10.1111/acfi.12851.
Prashant, A.S., Ghosh, J., 2023. The SEPs Debate and Surrounding Issues: Part -IV 28 (4). doi:10.56042/jipr.v28i4.2207.
Preble, J.F., 1997. Integrating the Crisis Management Perspective into the Strategic Management Process. Journal of Management Studies 34 (5), 769–791. doi:10.1111/1467-6486.00071.
Rassenfosse, G. de, Dernis, H., Guellec, D., Picci, L., van Pottelsberghe de Potterie, B., 2013. The worldwide count of priority patents: A new indicator of inventive activity. Research Policy 42 (3), 720–737. doi:10.1016/j.respol.2012.11.002.
Rassenfosse, G. de, Jaffe, A.B., 2018. Are patent fees effective at weeding out low-quality patents? Journal of Economics & Management Strategy 27 (1), 134–148. doi:10.1111/jems.12219.
Reitzig, M., 2003. What determines patent value? Research Policy 32 (1), 13–26. doi:10.1016/S0048-7333(01)00193-7.
Ribeiro, B., Shapira, P., 2020. Private and public values of innovation: A patent analysis of synthetic biology. Research Policy 49 (1), 103875. doi:10.1016/j.respol.2019.103875.
Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why Should I Trust You?", in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA. 13 08 2016 17 08 2016. ACM, New York, NY, USA, pp. 1135–1144.
Richter, K., Klos, M., 2022. Mehr Patentklagen in Deutschland – auch dank kreativer Münchner Richter (downloaded on 7 March 2023 from https://www.juve.de/markt-und-management/zahl-der-patentklagen-in-deutschland-steigt-deutlich/).
Righi, C., Simcoe, T., 2019. Patent examiner specialization. Research Policy 48 (1), 137–148. doi:10.1016/j.respol.2018.08.003.
Risch, J., Alder, N., Hewel, C., Krestel, R., 2020. PatentMatch: A Dataset for Matching Patent Claims & Prior Art http://arxiv.org/pdf/2012.13919v1).
Risch, J., Krestel, R., 2018. Learning Patent Speak: Investigating Domain-Specific Word Embeddings, in: 2018 Thirteenth International Conference on Digital Information Management (ICDIM). 2018 Thirteenth International Conference on Digital Information Management (ICDIM), Berlin, Germany. 24.09.2018 - 26.09.2018. IEEE, pp. 63–68.
Risch, J., Krestel, R., 2019. Domain-specific word embeddings for patent classification. Data Technologies and Applications 53 (1), 108–122. doi:10.1108/DTA-01-2019-0002.
Schmitt, V.J., 2024. Disentangling patent quality: Using a large language model for a systematic literature review. Submitted in Scientometrics.
Schmitt, V.J., Denter, N.M., 2024. Modeling an indicator for statutory patent novelty. World Patent Information 78, 102283. doi:10.1016/j.wpi.2024.102283.
Schmitt, V.J., Walter, L., Schnittker, F.C., 2023. Assessment of patentability by means of semantic patent analysis – A mathematical-logical approach. World Patent Information 73, 102182. doi:10.1016/j.wpi.2023.102182.

Schuett, F., 2013. Patent Quality and Incentives at the Patent Office. The RAND Journal of Economics 44 (2), 313–336.

Setchi, R., Spasić, I., Morgan, J., Harrison, C., Corken, R., 2021. Artificial intelligence for patent prior art searching. World Patent Information 64, 102021. doi:10.1016/j.wpi.2021.102021.

Shrestha, Y.R., Krishna, V., Krogh, G. von, 2021. Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. Journal of Business Research 123, 588–603. doi:10.1016/j.jbusres.2020.09.068.

Somaya, D., 2012. Patent Strategy and Management. Journal of Management 38 (4), 1084–1114. doi:10.1177/0149206312444447.

Spithoven, A., Vanhaverbeke, W., Roijakkers, N., 2013. Open innovation practices in SMEs and large enterprises. Small Business Economics 41 (3), 537–562. doi:10.1007/s11187-012-9453-9.

Squicciarinii, M., Dernisi, H., Criscuoloi, C., 2023. Measuring Patent Quality: Indicators of Technological and Economic Value 2023/05. doi:10.1787/18151965.

Tecuci, G., 2012. Artificial intelligence. WIREs Computational Statistics 4 (2), 168–180. doi:10.1002/wics.200.

The White House, 2022. The Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America: An economic study prepared in response to the US-EU Trade and Technology Council Inaugural Joint Statement, The United States Government (downloaded on 3 June 2024 from https://www.whitehouse.gov/wp-content/uploads/2022/12/TTC-EC-CEA-AI-Report-12052022-1.pdf).

Thiebes, S., Lins, S., Sunyaev, A., 2021. Trustworthy artificial intelligence. Electronic Markets 31 (2), 447–464. doi:10.1007/s12525-020-00441-4.

Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. The RAND Journal of Economics 21 (1), 172. doi:10.2307/2555502.

Trappey, A.J., Hsu, F.-C., Trappey, C.V., Lin, C.-I., 2006. Development of a patent document classification and search platform using a back-propagation network. Expert Systems with Applications 31 (4), 755–765. doi:10.1016/j.eswa.2006.01.013.

Trappey, A.J., Trappey, C.V., Wu, C.-Y., Lin, C.-W., 2012. A patent quality analysis for innovative technology and product development. Advanced Engineering Informatics 26 (1), 26–34. doi:10.1016/j.aei.2011.06.005.

Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Sun, J.J.H., 2021. Patent Value Analysis Using Deep Learning Models—The Case of IoT Technology Mining for the Manufacturing Industry. IEEE Transactions on Engineering Management 68 (5), 1334–1346. doi:10.1109/TEM.2019.2957842.

Tseng, Y.-H., Wu, Y.-J., 2008. A study of search tactics for patentability search, in: Proceeding of the 1st ACM workshop on Patent information retrieval - PaIR '08. Proceeding of the 1st ACM workshop, Napa Valley, California, USA. 30.10.2008 - 30.10.2008. ACM Press, New York, New York, USA, pp. 33–36.

Tzabbar, D., Silverman, B.S., Aharonson, B.S., 2015. Learning by hiring or hiring to avoid learning? Journal of Managerial Psychology 30 (5), 550–564. doi:10.1108/JMP-01-2013-0001.

USPTO, 2019a. 35 U.S.C. 102 Conditions for patentability; novelty. (downloaded on 21 December 2021 from https://www.uspto.gov/web/offices/pac/mpep/mpep-9015-appx-l.html#d0e302376).

USPTO, 2019b. 35 U.S.C. 101 Inventions patentable. (downloaded on 21 December 2021 from https://www.uspto.gov/web/offices/pac/mpep/mpep-9015-appx-l.html#d0e302376).

USPTO, 2019c. 35 U.S.C. 103 Conditions for patentability; non-obvious subject matter. (downloaded on 21 December 2021 from https://www.uspto.gov/web/offices/pac/mpep/mpep-9015-appx-l.html#d0e302376).

USPTO, 2019d. Manual of Patent Examining Procedure: 608.01(i) Claims [R-3] (downloaded on 213.02.2022 from https://mpep.uspto.gov/RDMS/MPEP/e8r9#/e8r9/d0e44872.html).

USPTO, 2019e. Manual of Patent Examining Procedure: Appendix L - Patent Laws (downloaded on 21 December 2021 from https://www.uspto.gov/web/offices/pac/mpep/mpep-9020-appx-r.html).

van Audenrode, M., Royer, J., Stitzing, R., SSSskilahti, P., 2017. Over-Declaration of Standard Essential Patents and Determinants of Essentiality. SSRN Electronic Journal. doi:10.2139/ssrn.2951617.

Varma, B.K., 2014. Intellectual Property Rights and the Technology Transfer Process, in: , Treatise on Process Metallurgy. Elsevier, pp. 1249–1289.

Venkatesh, Thong, Xu, 2012. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 36 (1), 157. doi:10.2307/41410412.

Venkatesh, V., 2000. Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Information Systems Research 11 (4), 342–365. doi:10.1287/isre.11.4.342.11872.

Verhoeven, D., Bakker, J., Veugelers, R., 2016. Measuring technological novelty with patent-based indicators. Research Policy 45 (3), 707–723. doi:10.1016/j.respol.2015.11.010.

Walter, L., Denter, N.M., Kebel, J., 2022. A review on digitalization trends in patent information databases and interrogation tools. World Patent Information 69, 102107. doi:10.1016/j.wpi.2022.102107.

Walter, L., Radauer, A., Moehrle, M.G., 2017. The beauty of brimstone butterfly: novelty of patents identified by near environment analysis based on text mining. Scientometrics 111 (1), 103–115. doi:10.1007/s11192-017-2267-4.

Walter, L., Schnittker, F.C., 2016. Patentmanagement: Recherche - Analyse - Strategie. De Gruyter Oldenbourg, Berlin.

Wang, B., Hsieh, C.-H., 2015. Measuring the value of patents with fuzzy multiple criteria decision making: insight into the practices of the Industrial Technology Research Institute. Technological Forecasting and Social Change 92, 263–275. doi:10.1016/j.techfore.2014.09.015.

Whalen, R., 2018. Boundary spanning innovation and the patent system: Interdisciplinary challenges for a specialized examination system. Research Policy 47 (7), 1334–1343. doi:10.1016/j.respol.2018.04.017.

White, S.K., 2023. What is a business analyst? A key role for business-IT efficiency (downloaded on 23 April 2024 from https://www.cio.com/article/276798/project-management-what-do-business-analysts-actually-do-for-software-implementation-projects.html).

Willim, H.-D., 2009. Verfahren zum Aufrichten eines Kranauslegers (downloaded on 3 March 2023 from [online] https://patents.google.com/patent/DE102007051539A1/de?oq=DE+102007051539B4+). Accessed 3 March 2023.

WIPO, 2019. WIPO Technology Trends 2019 - Artificial Intelligence: Executive summary (downloaded on 9 April 2024 from https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf).

WIPO, 2022. World Intellectual Property Report 2022: The Direction of Innovation (downloaded on 11 April 2024 from https://www.wipo.int/edocs/pubdocs/en/wipo-pub-944-2022-en-world-intellectual-property-report-2022-the-direction-of-innovation.pdf).

WIPO, 2023. World Intellectual Property Indicators 2023 https://www.wipo.int/edocs/pubdocs/en/wipo-pub-941-2023-en-world-intellectual-property-indicators-2023.pdf).

Wittfoth, S., 2019a. Identification of Probable Standard Essential Patents (SEPs) Based on Semantic Analysis of Patent Claims, in: 2019 Portland International Conference on Management of Engineering and Technology (PICMET). IEEE, pp. 1–9.

Wittfoth, S., 2019b. Measuring technological patent scope by semantic analysis of patent claims – An indicator for valuating patents. World Patent Information 58, 101906. doi:10.1016/J.WPI.2019.101906.

Wu, J.-L., Chang, P.-C., Tsao, C.-C., Fan, C.-Y., 2016. A patent quality analysis and classification system using self-organizing maps with support vector machine. Applied Soft Computing 41, 305–316. doi:10.1016/j.asoc.2016.01.020.

Wu, K., Zhao, Y., Zhu, Q., Tan, X., Zheng, H., 2011. A meta-analysis of the impact of trust on technology acceptance model: Investigation of moderating influence of subject and context type. International Journal of Information Management 31 (6), 572–581. doi:10.1016/j.ijinfomgt.2011.03.004.

Wu, Y., Ji, Y., Gu, F., Guo, J., 2021. A collaborative evaluation method of the quality of patent scientific and technological resources. World Patent Information 67, 102074. doi:10.1016/j.wpi.2021.102074.

Wustmans, M., 2019. Patent Intelligence zur unternehmensrelevanten Wissenserschließung. Springer Fachmedien Wiesbaden, Wiesbaden.

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. doi:10.1016/j.techfore.2020.120366.

Yang, S., 2023. Predictive Patentomics: Forecasting Innovation Success and Valuation with ChatGPT.

Yoon, J., Park, H., Kim, K., 2013. Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis. Scientometrics 94 (1), 313–331. doi:10.1007/s11192-012-0830-6.

Yun, J., Geum, Y., 2020. Automated classification of patents: A topic modeling approach. Computers & Industrial Engineering 147, 106636. doi:10.1016/j.cie.2020.106636.

Zhao, L., Xiang, Y., Yi, Q., 2017. Fuzzy front end patent management and innovation performance. Management Decision 55 (6), 1143–1162. doi:10.1108/MD-03-2016-0127.

Appendix A: Declaration of personal contribution

Table 4 shows the personal contribution of the author of this cumulative dissertation to the five considered publications ***P1 - P5*** with regard to proportion and type.

Table 4: Overview of personal contribution to the five considered publications.

	Authors	*Portion of contribution*	*Type of contribution*
P1	Schmitt (under review)	100%	Sole authorship
P2	Schmitt, Walter, Schnittker (2023)	60%	Conceptualization, Methodology, Software, Validation, Investigation, Writing – Original Draft, Writing - Review & Editing, Visualization.
P3	Walter, Schmitt (2023)	50%	Conceptualization, Methodology, Software, Validation, Investigation, Writing – Original Draft, Writing - Review & Editing, Visualization.
P4	Schmitt, Denter (2024)	60%	Conceptualization, Investigation, Methodology, Data curation, Software, Validation, Writing – Original Draft, Writing - Review & Editing, Visualization
P5	Herzberg, Schmitt, Walter (under review)	30%	Conceptualization, Investigation, Data curation, Software, Methodology, Visualization, Writing – original draft

Source: Author

Appendix B: Bibliographic information of publications *P1-P5*

The bibliographic information of the publications

- ***P1***: "*Disentangling patent quality: Using a large language model for a systematic literature review*",
- ***P2***: "*Assessment of patentability by means of semantic patent analysis – A mathematical-logical approach*",
- ***P3***: "*Detecting patent conflicts by means of computer-based feature analysis*" (Translated from German),
- ***P4***: "*Modeling an indicator for statutory patent novelty*",
- ***P5***: "*Semantic analysis for assessing the standard-essentiality of patents – Opportunities and challenges*",

are attached in the following.

P1: Disentangling patent quality: Using a large language model for a systematic literature review

Valentin J. Schmitt[1,*]

[1]Faculty of Business Studies and Economics, Institute of Project Management and Innovation, University of Bremen, Bremen, Germany

*Corresponding author

Submitted to *Scientometrics* journal:
Valentin J. Schmitt (tbd): Disentangling patent quality: Using a large language model for a systematic literature review. In *Scientometrics*.

Abstract

Assessing the quality or value of patents has long been the subject of research interest presenting various challenges such as interchangeable terminology, overlapping indicators, and diverse perspectives. Addressing these challenges, this study presents a comprehensive framework for assessing patent quality, that draws on stakeholder theory and adopts a multidimensional perspective encompassing economic, legal, and technological quality. This study identifies and categorizes quality indicators from the relevant literature as part of a systematic literature review using the large language model GPT-4. The results reveal that there is a predominant focus on multidimensional assessment of patent quality, particularly prioritizing technological quality, in almost two-thirds of the literature reviewed. The main findings suggest several research propositions, emphasizing the critical evaluation of indicators, the application of sophisticated methods, and the quantification of complex indicators. As a contribution to management, this study provides a comprehensive overview of existing patent quality indicators that enable a holistic assessment. The study contributes to scholarship by discussing the complexity of patent quality and by providing a framework for assessing patent quality from different perspectives. Furthermore, the use of large language models to improve the efficiency and effectiveness of systematic literature reviews is highlighted.

P2: Assessment of patentability by means of semantic patent analysis – a mathematical-logical approach

Valentin J. Schmitt[1,*], Lothar Walter[1], Frank C. Schnittker

[1]Faculty of Business Studies and Economics, Institute of Project Management and Innovation, University of Bremen, Bremen, Germany

*Corresponding author

Published in the *World Patent Information* journal:
Valentin J. Schmitt, Lothar Walter, Frank C. Schnittker (2023): Assessment of patentability by means of semantic patent analysis – a mathematical-logical approach. In *World Patent Information* 73. DOI: 10.1016/j.wpi.2023.102182.

Abstract

To obtain patent protection, a patent must fulfill statutory patentability requirements examined by a patent office. Such examinations are mostly performed manually and are quite time-consuming. Therefore, we suggest a computer-based process for the assessment of patentability by means of a mathematical-logical approach comparing patents with semantic structures. In order to make such an assessment, we compare the feature combinations of patent claims with the pertinent prior art. For proof of concept, the process has been tested successfully on an US-application claiming a method for raising a crane boom which can be categorized as non-patentable with regard to the requirement of non-obviousness. The result is consistent with that of a USPTO patent examiner, which underpins that at least under certain conditions not only patent examiners but also applicants and third parties can assess the chance and scope of protection for claimed inventions and patent applications with regard to patentability by our process.

P3: Detecting patent conflicts by means of computer-based feature analysis (Translated from German)

Valentin J. Schmitt[1,*], Lothar Walter[1]

[1]Faculty of Business Studies and Economics, Institute of Project Management and Innovation, University of Bremen, Bremen, Germany

*Corresponding author

Published in the Proceedings of the PATINFO 2023:
Valentin J. Schmitt, Lothar Walter, 2023. Aufdeckung von Schutzrechtskollisionen mittels computergestützter Merkmalsanalysen, in: PATINFO2023 "Schutzrechtsinformationen als Rohstoff für die Wirtschaft von morgen". 45. Kolloquium der TU Ilmenau über Patentinformation und Gewerblichen Rechtsschutz. Technische Uni Ilmenau, Ilmenau, Thür, pp. 163–177.

Abstract

Patente sind Ausschließlichkeitsrechte, die den Patentinhaber berechtigen den Schutzgegenstand zu nutzen und anderen die Benutzung zu verbieten. Auch wenn von Amts wegen der Schutzgegenstand beurteilt und hinsichtlich der Patentfähigkeit geprüft wird, kommen in der Praxis immer wieder Schutzrechtskollisionen vor, die beispielsweise zu einem Patentnichtigkeitsverfahren führen können. Die Prüfung, ob eine Schutzrechtskollision vorliegt, erfolgt anhand einer Merkmalsanalyse von Patentansprüchen. Hierzu werden die einzelnen Merkmale, welche die unter Schutz gestellte technische Neuerung einer Erfindung darstellen, in einer tabellarischen Zusammenstellung aufgegliedert und analysiert. Da für eine solche Merkmalsanalyse eine Vielzahl an Patenten verglichen werden, ist eine Merkmalsanalyse mit einem hohen zeitlichen und manuellen Aufwand verbunden. Um diesen Aufwand zu minimieren, stellen wir in diesem Beitrag eine computergestützte Merkmalsanalyse vor, die durch den Einsatz von Textmining-Techniken auch den Wortlaut der zu analysierenden Patentansprüche berücksichtigt. Anhand einer beim Deutschen Patent- und Markenamt zur Patentierung eingereichten Erfindung zum Anheben eines Kranauslegers vergleichen wir computergestützt die Merkmalskombinationen der Patentansprüche mit dem

einschlägigen Stand der Technik und validieren die mit der vorgestellten Merkmalsanalyse aufgedeckten Schutzrechtskollisionen.

P4: Modeling an indicator for statutory patent novelty

Valentin J. Schmitt[1,*], Nils M. Denter[1]

[1]Faculty of Business Studies and Economics, Institute of Project Management and Innovation, University of Bremen, Bremen, Germany

*Corresponding author

Published in the *World Patent Information* journal:

Valentin J. Schmitt, Nils M. Denter (2024): Modeling an indicator for statutory patent novelty. In *World Patent Information* 78. DOI: 10.1016/j.wpi.2024.102283.

Abstract

Novelty is considered a conditio sine qua non for the grant of a patent by most relevant patent authorities and in U.S. patent law defined by 35 U.S.C. §102. Previous attempts to operationalize patent novelty have been mostly based on theoretical principles, such as recombination theory, and have not estimated novelty according to data that officially determines whether sufficient novelty is present in an application. To overcome this gap, this study analyzes whether established measures of patent novelty are capable of predicting the rejection of an application based on lack of novelty. Furthermore, this study applies a combination of sophisticated unsupervised and supervised machine learning techniques to patent data and provides the possibility to estimate statutory novelty practiced by the USPTO by a modeled indicator. Measuring such statutory novelty would give applicants a tremendous competitive advantage to pursue patent strategies offensively and/or defensively. For example, the indicator allows companies – particularly small and medium-sized companies with limited resources – to assess the novelty and therefore the patentability of one's own invention. Large companies, most of which are more likely to pursue an offensive patent strategy, can use the indicator to measure the novelty of numerous published third-party patents and challenge their validity.

P5: Semantic analysis for assessing the standard-essentiality of patents – Opportunities and challenges

Andre Herzberg[1,*], Valentin J. Schmitt[1], Lothar Walter

[1]Faculty of Business Studies and Economics, Institute of Project Management and Innovation, University of Bremen, Bremen, Germany

*Corresponding author

Submitted to *IEEE Transactions on Engineering Management* journal:
Andre Herzberg, Valentin J. Schmitt, Lothar Walter (tbd): Semantic analysis for standard-essentiality of patents – Opportunities and challenges. In *IEEE Transactions on Engineering Management*.

Abstract

Interoperability between technical devices is crucial for manufacturers as well as consumers and is largely facilitated by the development and adoption of technical standards. These standards, which are prominently utilized in Information and Communication Technology, are developed by standard setting organizations such as the European Telecommunications Standards Institute. Companies involved in this process can influence the standards to secure competitive advantages, often by declaring their patents as standard-essential. However, the lack of a rigorous evaluation of these declarations can lead to issues such as over-declaration or under-declaration of standard-essential patents (SEPs). Determining the SEP status is complex, involves technical and sometimes economic considerations, and is typically a time-consuming manual process. Recent research has explored semantic and machine learning methods for automating SEP assessment, although these methods primarily provide probabilistic outcomes. Our current research aims to improve existing semantic analysis approaches by addressing the distinct linguistic and structural features of technical standards and patents. The results indicate which type of pre-processing and which text fragment from a selected technical standard and US-patents yields in the highest semantic similarity. This has methodological implications for researchers, as its sheds light on the assessment of standard-essentiality. Furthermore, managerial implications arise for companies, policy makers and patent examiners, who can use our approach for the assessment of SEPs and as yet undeclared patents