GATR Journal of Business and Economics Review

GATR JOURNALS

Journal homepage: http://gatrenterprise.com/GATRJournals/JBER/vol10_2025_issue3.html

J. Bus. Econ. Review 10(3) 131 – 142 (2025)

The Elaboration of the Patent Processing Instrument Based on Machine Learning Technology

Cross Mark

Sheresheva, M.Y.[1]*, Gorlacheva, E.N.[2]

[1]Lomonosov Moscow State University, Leninskie Gory 1-46, 119991, Moscow, Russia
[2]Bauman Moscow State Technical University, 2nd Baumanskaya st. 5, 105005, Moscow, Russia

ABSTRACT

Objective – While managing the innovation activity, it is necessary to base it on reliable sources of scientific and technical information, including patent research. However, the existing variety and scale of patent databases necessitate the development of an instrument that enables processing large volumes of patent information within limited timeframes. In these conditions, it is necessary to use machine learning (ML) technology to create a solid information base for management decisions.
Methodology – The objective of the study presented in the paper was to propose an algorithm for processing patent data to improve the quality of patent research. The essence of the algorithm is that all necessary patents are ranked according to a relevance criterion, after which the researcher analyzes the already essential patents.
Findings – The paper envisages the algorithm's practical realization using a gravity-driven power generator case. Findings indicate that the proposed new instrument enables a significant reduction in processing time for patent data.
Novelty – The paper contributes to innovation management by integrating patent analytics and machine learning.
Type of Paper: Empirical
JEL Classification: D80, D81
Keywords: Innovation activity; patent analytics; machine learning technology; a gravity-driven power generator
Reference to this paper should be made as follows: Sheresheva, M.Y.; Gorlacheva, E.N. (2025). The elaboration of the patent processing instrument based on the machine learning technology, *J. Bus. Econ. Review,* 10(3), 131–142. https://doi.org/10.35609/jber.2025.10.3(5)

1. Introduction

Innovation activity is one of the main drivers of development in the cyber economy. The search for new technologies and business solutions creates increased demand for the results of intellectual activity, which are a key asset of an enterprise. An essential source for management decisions in innovation activity is patent analytics, as it reflects not only the level of current research and development but also serves as a forecasting source for the operational and long-term directions of an industrial enterprise's innovative development. The researchers in the patent analysis sphere evaluate the place of patent analytics differently in innovation management (Moehrle et al., 2010) (Nikolaev, 2019).

* Paper Info: Revised: October 21, 2025
Accepted: December 31, 2025
* Corresponding author: Sheresheva, M.Y.
E-mail: m.sheresheva@mail.ru
Affiliation: Lomonosov Moscow State University, Leninskie Gory 1-46, 119991, Moscow, Russia

The relative novelty of the research field that forms at the junction of statistics, big data, economy, and innovatics, and the lack of the corresponding toolkit hampers the complex study of this direction both in theoretical and practical domains (Moehrle et al., 2010). However, the task of finding promising technological directions turns into an information overload for employees who do not always correctly cope with the tasks set. This often leads to mistakes in finding promising technological directions that can result in billions of dollars in losses or bankruptcy for companies (Daghfous et al., 2023) (Özsungur et al., 2023). There is a scientific and practical need to find tools that improve the "search" and "selection" stages by automating the analysis of the technological landscape. Traditional methods – expert surveys and patent analysis – do not enable timely technological forecasting. The integration of ML tools into the process of technological forecasting will improve the accuracy of identifying emerging technologies (Gurcan et al., 2023). The main advantage of using ML tools over traditional ones is speed, scale, and objectivity (Ahmed et al., 2023).

According to (Chen & Chang, 2012)There has been a significant development in patent analytics over the last two decades. Due to the digitalization of patent data, accessibility costs have decreased immensely. The rise of artificial intelligence, machine learning, deep learning, and artificial neural networks has provided a number of methods for analyzing massive data (Rahmani et al., 2021) (Xu et al., 2021). However, these technologies are hardly applied in patent research.

The purpose of the research is to close this gap by developing a machine-learning (ML)- based instrument for patent processing. ML algorithms allow high-tech enterprises to better "shape" technological capabilities than their competitors (Madanchian & Taherdoost, 2024) (Yavuz & Çalik, 2025).

The study presented in the paper is based on the theoretical framework of the resource-based view and the absorptive capacity concept, enabling the analysis of the contribution of ML tools to decision-making processes in R&D management.

The paper structure is as follows: Section 2 discusses the role of patent analytics in innovation and knowledge management, as well as the prospects of ML-based patent analytics for decision-making in R&D strategy. Section 3 describes the chosen methodology appropriate for this study. Section 4 briefly discusses the main results of the survey: reducing the volume of patents for analytical work, decreasing the time required for patent analysis, and improving the quality of patent searches due to the implementation of the elaborated ML algorithm. Finally, Section 5 concludes by outlining study limitations and proposing paths for future research.

2. Literature Review

2.1Theoretical concepts of innovation management

There are several well-known and noteworthy concepts in innovation management: resource-based view theory (Barney, 2001) and the concept of companies' absorptive capabilities (Fabrizio et al., 2022). Within the framework of the RBV, the key idea is that an enterprise achieves sustainable competitive advantages by possessing unique, valuable assets, rare and irreproducible resources, and capabilities. For companies that are actively engaged in innovation, patent data is considered a strategic asset (Bagna et al., 2021) (Holgersson & Granstrand, 2022).

The digital economy enhance information accessibility, drives corporate R&D innovation via direct and indirect mechanisms, facilitates open innovation, fosters the development of innovation ecosystems with strong network externalities, and promotes organizational change (Costa & Matias, 2020) (Kuzior et al.,

2023) (Mbatha, 2025) (Shao & Wang, 2025) (Zhao & Wang, 2025). Strong integration of knowledge management and innovation management becomes an essential source of competitive advantage; Ávila, 2022; Rehman et al., 2022; Due to new digital technologies, "various forms of knowledge-based assets may be absorbed, assimilated, shared and utilized for innovation" (Goh, 2005).

In particular, one can consider ML algorithms capable of extracting useful information from patents as dynamic capabilities of enterprises. The concept of absorptive capacity suggests that the availability of analytical tools allows companies to respond more clearly to emerging changes in the technological landscape, recognize the value of new, external information, integrate it into their activities, and apply it for commercial purposes (Kastelli et al., 2024) (Tallarico et al., 2025). The use of ML tools also makes it possible to cope with the "paradox of abundance" - the redundancy of unstructured external knowledge, translating raw patent data into structured, interpretable knowledge ready for transformation, identifying semantic relationships, problems, and trends. Combining these two concepts makes it possible to consider ML tools as a dynamic ability of a company to enhance technological forecasting. ML tools are becoming an organizational mechanism for solving innovation management problems.

2.2 The patent analytic process

Patent analytics can be defined as the science of analyzing large amounts of patent information to derive meaningful insights to support decision-making, which involves the use of different technologies, techniques and approaches (Aristodemou et al., 2017). In other words, it is a set of search, processing and analysis methods, aimed at applying new technologies in patent research.

The founders of the patent analytics idea can be considered the large research group under the auspices of the World Intellectual Property Organization (WIPO), led by A. Trippe (A. J. Trippe, 2003). Patent analytics describes techniques for analyzing large sets of patents in relation to other data sources while identifying relationships and trends (Bogdanova et al., 2019). Thanks to digitalization, patent data is becoming increasingly accessible to non-specialists (Aristodemou & Tietze, 2017).

Concerning the theoretical domain, the patent analytics knowledge corpus is mainly developing in two interconnected directions: the instrumental approach and the conceptual approach. The first one deals exclusively with means of working with patent data. Examples of the instrumental approach are specialized software products in the field of patent data analysis and processing (Moehrle et al., 2010). This approach is supported, namely, by practitioners and developers of specialized software. The conceptual approach to patent analytics is a rapidly developing area of scientific research aimed at improving the theoretical foundations of patent data analysis (Chuprat et al., 2024). Thus, within the framework of this approach, specialists develop methods for determining the strength and significance of a patent, the subject area of research, or the division of patent collections for separate analysis of patent documents.

In accordance with (Abbas et al., 2014) (A. Trippe, 2015) (Aristodemou & Tietze, 2018)The patent analytics comprises three main stages: pre-processing, processing, and post-processing (Fig. 1). In the pre-processing stage, the data is collected, cleaned, prepared comprehensively, and the information extracted. During the processing stage, the extracted data are analyzed and classified using various methods. The post-processing stage is known as knowledge discovery, where the results are visualized and evaluated, thereby supporting strategic decision-making.

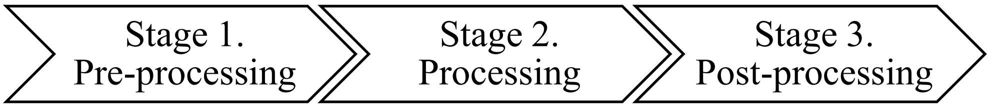

Figure 1. The patent analytics process.

With the ever-increasing volumes of patent information, the tasks of patent research and analysis have become vital from the managerial perspective (Park et al., 2013) (J. Choi et al., 2021) (Liu et al., 2021). Patent documents contain large volumes of structured and unstructured data that require advanced, intelligent techniques for analysis. Traditional patent analysis involves a lot of effort and time. The use of ML technology in the field helps substantially reduce costs.

2.3 International trends in the patent activity

The intellectual property constitutes the kernel of the modern global economy – the appearance of the new results of the intellectual creative activities; its legal protection precedes the material movement of goods and services (Hausmann et al., 2024).

In the near future, experts expect an increase in the digital network traffic connected to the transfer of intellectual property objects intended for industrial production (Hübscher et al., 2022) (Lu, 2022). Robotic manufacturing and additive technologies development serve as catalyst for this process (Sekerin et al., 2021) (Bhatia et al., 2024). This allowed active use of digital three-dimensional images, as well as digital descriptions of industrial objects, in production (Wen et al., 2021) (Shang et al., 2022).

WIPO experts note the positive dynamics of global patent activity. In 2018, a new patenting record was set - 3.3 million applications for patents for inventions, as well as 1.3 million industrial designs and 2.1 million utility models (WIPO, 2019) (Di et al., 2022) showed that over the last decades, the global innovation network has expanded rapidly from a sparse to a dense, complex network. What is also worth noting is that the global innovation network shows an evident "core-periphery" pattern (Kemeny et al., 2025). Patent application is unevenly distributed around the globe. Countries such as the US, China, and Canada have been the top countries flowing in, while Japan, Korea, the EU, and Switzerland have been the central countries flowing out (Di et al., 2022).

2.4 The role and the place of patent analytics in innovation management

While managing innovation, patent analytics contributes to the solution of several tasks. First, it allows us to find and select the most promising and effective scientific and technical solutions that lead to the creation of appropriate technologies and equipment in the future (Yun et al., 2022) (Yang et al., 2023). Second, it makes it possible to determine trends in the development of technologies (Y. Choi et al., 2021) (J. Wang & Hsu, 2021). In addition, patent analytics allows for assessing the competitiveness and technical level of manufactured or developed industrial products, and to analyze the market situation to identify the main competitors and possible partners (Glushko & Silnov, 2021) (Lee, 2021) (Shuai et al., 2022).

The information base for patent research comprises an extensive array of public-domain sources of patent, scientific, and technical information. The openness of these patent documents and ensuring the completeness of the disclosure of technical details are the most important advantages of patent research (Ryabokon et al., 2019) (Dyer et al., 2024).

The essential factors that make possible successful use of patent bases for analysis are as follows: priority date of patent documents; similar requirements for the volume and form of the claimed technical solutions presentation in patent documents of different countries; patent documents structure in accordance with the International Patent Classification; mandatory indication of the author(s) and patent owner(s); links to previously known technical solutions and similar patents.

In the digital era, providing reliable and effective patent research at the enterprise should be carried out at the expense of the qualified personnel who have the skills to work with modern software tools for patent analytics (Szczepańska-Woszczyna & Gatnar, 2022) (Aithal & Aithal, 2023). Machine learning is an important part of the contemporary toolkit that allows improvement of the patent analysis process.

2.5 Machine learning technologies in processing the patent data

Machine learning (ML) is a set of mathematical, statistical, and computational methods for developing algorithms that can solve problems indirectly by identifying patterns in a variety of input data. This is the practice of using algorithms to analyze data, study it, and then determine or predict something (Gorlacheva & Goncharova, 2020) (Gorlacheva & Goncharova, 2021).

(C.-C. Wang & Chien, 2025) studied the transformative impacts of ML technologies on Taiwan's high-precision, high-mix manufacturing industries and confirmed that industry champions such as Delta Electronics, Foxconn, and TSMC have made measurable improvements in process performance by employing CNNs for defect detection, reinforcement learning for scheduling, and federated learning for cross-factory optimization. Several recent studies have shown that ML technologies can reduce the burden on the departments that make management decisions (Goodfellow et al., 2016) (Ali et al., 2023) (Taherdoost, 2023).

There is also evidence that ML techniques, including classification, regression, and clustering, have emerged as powerful tools for predicting patent strength. They leverage patterns and relationships within large-scale patent datasets, enabling automated and data-driven approaches (Bhatt et al., 2024). With the help of ML algorithms, you can clean up data, eliminate gaps in it, predict, rank, classify, identify anomalies and outliers, etc. (Nassif et al., 2021) (Borrohou et al., 2023).

3. Research Methodology

The research methodology involved several stages: collecting data from various patent databases; cleaning and normalizing the text (in the example given, patents were in different languages and had to be translated into English); using the semantic distance method (cosine similarity); and a ranking model. The quality of the algorithm was evaluated using a normalized DCG based on query relevance.

The patent data is a formal information source containing codified information, digital data, and text data. Therefore, for machine analysis of patent data, it is necessary to use various ML methods tailored to each data type. The complexity of this approach lies in the differences in the forms of patent documents in different countries.

The use of machine learning technologies in patent analytics reduces time and human resources and increases the quality of results. The sources of data are various: Espacenet, Patentoscope, EAPO.ORG, EAPATIS.COM, USPTO, Patent Yandex, Google Patent, etc. In this paper, we use the European Patent Office's Espacenet database. As part of the patent search, there was a gravitational electric generator. The starting query for the database was the following:

pd > "2015" AND ctxt all "gravity electric generator" AND ctxt all "gravity power generator."

For this query, 1,775 results have been received. During the analysis of the results, many documents were in different languages. As searching and analyzing different languages is quite complex, the English language has been selected as the primary language. All the patents were automatically translated on it. Using the elaborated instrument, the data were collected, transformed into the required format, and saved as JSON.

To rank the selected patents, we used the "Learning to rank" method, which is applied in the field of Big Data search and ranking. The simple ranking model is supposed to build a vector space between the patent description and our queries (1):

$$sim(d,q) = cos(\vec{v}(d),\vec{v}(q)) = \frac{\vec{v}(d)\cdot\vec{v}(q)}{||\vec{v}(d)||\cdot||\vec{v}(q)||} = \frac{\sum_{i=1}^{|V|} d_i \cdot q_i}{\sqrt{\sum_{i=1}^{|V|} d_i^2}\cdot\sqrt{\sum_{i=1}^{|V|} q_i^2}} \quad (1)$$

```
# Connecting Packages
import numpy as np
import pandas as pd
import re
import math

# Read data from json file
open_data = pd.read_json('main_data.json', orient='index')
open_data.index.name = "ID"

# Creating search queries
objects_to_search_one = ['gravity', 'electric', 'generator']
objects_to_search_two = ['gravity', 'power', 'generator']

# Merge queries into a dictionary of terms
objects_to_search = set(objects_to_search_one).union(set(objects_to_search_two))

# Removing patents without a description
open_data_main = open_data.dropna(subset=['Abstract'])

# Take only the patent description from the data and calculate the terms frequency
search_df = pd.DataFrame(open_data_main["Abstract"])

for word in objects_to_search:
    search_df[word] = open_data_main["Abstract"].str.count(word,flags=re.IGNORECASE)
```

Based on the results of this code snippet, the frequency of use of search query terms in the patent description was obtained. By deducing the total frequency of use of the terms, it was revealed that the terms "gravity" and "generator" are unique.

```
# The term gravity and generate is unique. Delete all patents that do not contain them in the description.
drop_arr = []
for i in range(len(open_data_main)):
    if not "gravity" in open_data_main.iloc[i]['Abstract']:
```

```
            drop_arr.append(open_data_main.index[i])

    open_data_main = open_data_main.drop(index=drop_arr)
    search_df = search_df.drop(index=drop_arr)

    drop_arr = []
    for i in range(len(open_data_main)):
        if not "generator" in open_data_main.iloc[i]['Abstract'] and not "generator" in open_data_main.iloc[i]['Patent Title']:
            drop_arr.append(open_data_main.index[i])

    open_data_main = open_data_main.drop(index=drop_arr)
    search_df = search_df.drop(index=drop_arr)
```

After clearing the data from patents that do not contain the unique terms "gravity" and "generator", a sample of 393 patents was obtained.

```
# Creating the main functions for calculating the distance between vectors
def get_vec_d(doc):
    doc_arr = doc.lower().split(" ")
    frequency = {}
    for word in doc_arr:
        count = frequency.get(word,0)
        frequency[word] = count + 1

    len_freq = len(frequency)
    f_list = frequency.keys()

    sqrt_word = 0
    for word in f_list:
        if frequency[word] == 1:
            len_freq -= 1
        else:
            sqrt_word += frequency[word] ** 2

    ret_res = (sqrt_word + len_freq) ** 0.5

    return ret_res

def get_cos_d_q(data, query):
    data_len = len(data)
    ret_arr = []
    for i in range(data_len):
        numer = 1
        for word in query:
            numer += data[word].iloc[i]

        denom = get_vec_d(data["Abstract"].iloc[i]) * (len(query) ** 0.5)
```

```
        if denom == 0:
            ret_arr.append(0)
        else:
            cos_d_q = numer/denom
            ret_arr.append(cos_d_q)

    return ret_arr

# Calculating the distance between query and description vectors
search_df['Query 1'] = get_cos_d_q(search_df, objects_to_search_one)
search_df['Query 2'] = get_cos_d_q(search_df, objects_to_search_two)

# Merging the main data and the obtained
main_data = pd.merge(open_data_main, search_df, how='inner', on='Abstract')

# Ranking all data
main_data.sort_values(by = ['Query 2', 'Query 1'], ascending = False)
```

4. Results and Discussion

After calculating the distances between the query vectors and the patents, and applying ranking based on the results, the following sequence of patents by relevance was obtained. The elaborated instrument enabled sorting 1,775 patents by the keyword "gravity electric generator". After aggregating the use frequency of patent terms, we found that the terms “gravitation” and “generator” are unique. After cleaning the data from patent documents that don’t contain these two unique terms, a sample of 393 patents has been received that should be processed manually by an analyst in the future. Thus, the quality of patent research improved: the algorithm reduced the analyst's labor intensity by 4.5 times. Consequently, the time required to conduct patent research is reduced.

Russian enterprises most often assess the quality of patent research by the number of patents processed. The proposed algorithm helps rank patents by relevance, thereby improving the quality of patent research. The researcher, analyzing the document manually, initially examines a copy already suitable for the research topic. Thus, the introduction of machine learning allowed us to reduce the volume of patents for a patent search specialist's analytical work, reduce the time required for patent analysis, and improve the quality of patent search.

It should be noted that the implementation of this algorithm, when ranking data in different fields of knowledge and training a neural network, allows users to quickly find and analyze high-quality information clearly on request. In the future, if there is user feedback, the quality of the DCG algorithm can be assessed. Based on this, it is possible to determine patent research quality within the framework of the technology readiness assessment model, not only by the number of patents processed, but also using the DCG assessment.

The results obtained indicate the potential for patent analysts to use ML tools, as these tools allow us to "capture" unobvious (hidden) semantic links between patents that are inaccessible when searching only by keywords. Using the ranking model allows us not only to determine the binary relevance of patents, but also to give a preliminary assessment of their potential significance. Automation of routine patent search tasks also reduces the complexity of the initial screening of irrelevant documents.

The key advantage of ML algorithms is their ability to handle large amounts of data. The tool helps monitor active competitors, identifying unoccupied "white spots" in the technological space, and prioritizing areas for their own research in innovation management. ML algorithms are also applicable to consulting related to intellectual property, as they can significantly speed up the preparation of patent landscape reviews, reducing costs and increasing the throughput of patent analysts. At the strategic level, ML algorithms can help to map national technological competencies and evaluate the effectiveness of government R&D support programs to form priorities for scientific and technological development.

Of course, despite their advantages, ML algorithms have limits in their applicability. Their effectiveness has so far been tested in a well-structured technological field with well-established terminology.

5. Conclusion

In the context of digital transformation, there is a need to develop new approaches to managing enterprises' innovative activities. Patent data is essential for a variety of purposes, from identifying technological competitors to forecasting technological developments in a particular domain. Patent analytics can not only improve the quality of management decisions but also ensure their transparency, thereby avoiding an incorrect innovative development strategy.

Within the framework of the study, the theoretical foundations of patent analytics were analyzed, and the practical realization of an ML algorithm was presented in the context of a gravity-driven power generator. The study results confirmed that ML algorithms can effectively extract useful information from patents and therefore serve as dynamic capabilities of organizations engaged in innovation activities. The implementation of the elaborated ML algorithm has reduced the number of patents for the analytical work, sharply reducing the time required for patent analysis, and improving the quality of patent search. Therefore, one can conclude that patent analysis instruments would be worthwhile for managers engaged in strategic technology planning.

Despite the impressive results, the developed tool has several limitations. Firstly, for ML algorithms to become part of a patent analyst's daily routine, they need to have the appropriate qualifications of a data scientist. Secondly, the use of ML algorithms requires preliminary data preparation (purification, normalization, etc.). At the same time, the method can be helpful in a well-structured technological area with well-established terminology. If we are talking about a poorly formalized query, the algorithms may not show such impressive results. Thirdly, in case of errors or retraining of the model, the patent analyst will have to deal with technical issues – debugging the model rather than with the substantive part. How much such a tool will speed up his work is an ambiguous question.

In the future, it is necessary to test the proposed algorithm across various areas of patent search, particularly those with poorly structured data. With the development of generative models, the issue of using ML algorithms is also ambiguous. It is likely that LLM-based agents will cope much more efficiently and will not require additional training of patent analysis specialists. So far, one can state that in the context of the "redundancy paradox" of unstructured patent information, the use of classical patent search methods is clearly insufficient. Further research is essential to identify the most promising ways to implement the proposed algorithm and the most efficient digital tools for patent processing.

Acknowledgements

The Russian Science Foundation supports the study grant №25-18-00075, https://rscf.ru/project/25-18-00075/

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.

Ahmed, S. F., Alam, M. S. Bin, Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Shawkat Ali, A. B. M., & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521–13617.

Aithal, P. S., & Aithal, S. (2023). Introducing systematic patent analysis as an innovative pedagogy tool/experiential learning project in HE Institutes and Universities to boost awareness of patent-based IPR. *International Journal of Management, Technology, and Social Sciences (IJMTS)*, *8*(3), 395–413.

Ali, R., Hussain, A., Nazir, S., Khan, S., & Khan, H. U. (2023). Intelligent decision support systems—An analysis of machine learning and multicriteria decision-making methods. *Applied Sciences*, *13*(22), 12426.

Aristodemou, L., & Tietze, F. (2017). *A literature review on the state-of-the-art on intellectual property analytics*.

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.

Aristodemou, L., Tietze, F., Athanassopoulou, N., & Minshall, T. (2017). *Exploring the future of patent analytics: A technology roadmapping approach*.

Ávila, M. M. (2022). Competitive advantage and knowledge absorptive capacity: The mediating role of innovative capability. *Journal of the Knowledge Economy*, *13*(1), 185–210.

Bagna, E., Cotta Ramusino, E., & Denicolai, S. (2021). Innovation through patents and intangible assets: effects on growth and profitability of European companies. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(4), 220.

Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management*, *27*(6), 643–650.

Bhatia, V., Kumar, A., Sidharth, S., Khare, S. K., Ghorpade, S. C., Kumar, P., & AlZohbi, G. (2024). Industry 4.0 in Aircraft Manufacturing: Innovative Use Cases and Patent Landscape. In *Industry 4.0 Driven Manufacturing Technologies* (pp. 103–137). Springer.

Bhatt, N., Tripathi, A., Saini, D., Mehta, S., & Vijh, R. (2024). Patent Strength Predictor Using ML and Mamdani-Based FIS. *Biennial International Conference on Future Learning Aspects of Mechanical Engineering*, 229–241.

Bogdanova, E. L., Brovka, G. M., Maximova, T. G., & Nikolaev, A. S. (2019). Digital culture, skills of innovative entrepreneurship and intellectual property management: upcoming competences. *Innovations*, *252*(10), 101–109.

Borrohou, S., Fissoune, R., & Badir, H. (2023). Data cleaning survey and challenges–improving outlier detection algorithm in machine learning. *Journal of Smart Cities and Society*, *2*(3), 125–140.

Chen, Y.-L., & Chang, Y.-C. (2012). A three-phase method for patent classification. *Information Processing & Management*, *48*(6), 1017–1030.

Choi, J., Lee, J., & Yoon, J. (2021). Anticipating promising services under technology capability for new product-service system strategies: An integrated use of patents and trademarks. *Computers in Industry*, *133*, 103542.

Choi, Y., Park, S., & Lee, S. (2021). Identifying emerging technologies to envision a future innovation ecosystem: A machine learning approach to patent data. *Scientometrics*, *126*(7), 5431–5476.

Chuprat, S., Novianto, E. H. D., Matsuura, Y., Mahdzir, A. M., & Harun, A. N. (2024). A closer look on patent analytics through systematic literature review. *Management Review Quarterly*, 1–31.

Costa, J., & Matias, J. C. O. (2020). Open innovation 4.0 as an enhancer of sustainable innovation ecosystems. *Sustainability*, *12*(19), 8112.

Daghfous, A., Amer, N. T., Belkhodja, O., Angell, L. C., & Zoubi, T. (2023). Managing knowledge loss: a systematic literature review and future research directions. *Journal of Enterprise Information Management*, *36*(4), 1008–1031.

Di, Y., Zhou, Y., Zhang, L., Indraprahasta, G. S., & Cao, J. (2022). Spatial pattern and evolution of global innovation network from 2000 to 2019: global patent dataset perspective. *Complexity*, *2022*(1), 5912696.

Dyer, T. A., Glaeser, S., Lang, M. H., & Sprecher, C. (2024). The effect of patent disclosure quality on innovation. *Journal of Accounting and Economics*, *77*(2–3), 101647.

Fabrizio, C. M., Kaczam, F., de Moura, G. L., da Silva, L. S. C. V., da Silva, W. V., & da Veiga, C. P. (2022). Competitive advantage and dynamic capability in small and medium-sized enterprises: a systematic literature review and future research directions. *Review of Managerial Science*, *16*(3), 617–648.

Glushko, A., & Silnov, D. S. (2021). Analytics in the Field of Patent Documents. *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 364–368.
Goh, A. L. S. (2005). Harnessing knowledge for innovation: an integrated management framework. *Journal of Knowledge Management*, *9*(4), 6–18.
Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, Issue 2). MIT press Cambridge.
Gorlacheva, E. N., & Goncharova, N. P. (2020). Review of intellectual methods of management of innovative scientific and technical projects. *Nanotechnologies: Development, Application-XXI Century*, *14*(4), 5–19.
Gorlacheva, E. N., & Goncharova, N. P. (2021). Machine learning methods in project management. *In 10th Charnovsky Readings, Proceedings of the X All-Russian Scientific Conference on the Organization of Production*, 25–31.
Gurcan, F., Ayaz, A., Menekse Dalveren, G. G., & Derawi, M. (2023). Business intelligence strategies, best practices, and latest trends: Analysis of scientometric data from 2003 to 2023 using machine learning. *Sustainability*, *15*(13), 9854.
Hausmann, R., Yildirim, M. A., Chacua, C., Hartog, M., & Matha, S. G. (2024). Global trends in innovation patterns: A complexity approach. *World Intellectual Property Organization (WIPO) Economic Research Working Paper Series*, *80*.
Holgersson, M., & Granstrand, O. (2022). Value capture in open innovation markets: the role of patent rights for innovation appropriation. *European Journal of Innovation Management*, *25*(6), 320–339.
Hübscher, G., Geist, V., Auer, D., Ekelhart, A., Mayer, R., Nadschläger, S., & Küng, J. (2022). Graph-based managing and mining of processes and data in the domain of intellectual property. *Information Systems*, *106*, 101844.
Kastelli, I., Dimas, P., Stamopoulos, D., & Tsakanikas, A. (2024). Linking digital capacity to innovation performance: The mediating role of absorptive capacity. *Journal of the Knowledge Economy*, *15*(1), 238–272.
Kemeny, T., Petralia, S., & Storper, M. (2025). Disruptive innovation and spatial inequality. *Regional Studies*, *59*(1), 2076824.
Kuzior, A., Sira, M., & Brożek, P. (2023). New technologies as tools supporting innovation management. *Scientific Papers of Silesian University of Technology: Organization and Management Series*, *188*, 279–297.
Lee, P.-C. (2021). Investigating long-term technological competitiveness: Originality, generality, and longevity. *IEEE Transactions on Engineering Management*, *71*, 20–42.
Liu, N., Shapira, P., Yue, X., & Guan, J. (2021). Mapping technological innovation dynamics in artificial intelligence domains: Evidence from a global patent analysis. *Plos One*, *16*(12), e0262050.
Lu, J. (2022). Challenges and Responses to Intellectual Property Protection in New Industries. *Science of Law Journal*, *1*(3), 33–38.
Madanchian, M., & Taherdoost, H. (2024). AI-Powered Innovations in High-Tech Research and Development: From Theory to Practice. *Computers, Materials & Continua*, *81*(2), 2133–2159.
Mbatha, S. (2025). Sectoral systems of innovation in two BRICS countries: A case of the clothing, textile, leather, and footwear sector of South Africa and Brazil. *BRICS Journal of Economics*, *6*(2), 117–138.
Migdadi, M. M. (2021). Knowledge management, customer relationship management and innovation capabilities. *Journal of Business & Industrial Marketing*, *36*(1), 111–124.
Moehrle, M. G., Walter, L., Bergmann, I., Bobe, S., & Skrzipale, S. (2010). Patinformatics as a business process: A guideline through patent research tasks and tools. *World Patent Information*, *32*(4), 291–299.
Mohammed, R. K., Ahmed, N. E., Aziz, E. M., & Dewi, A. (2023). Knowledge management plays a crucial role in attaining a competitive advantage. *Al-Idarah: Jurnal Kependidikan Islam*, *13*(2), 113–123.
Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine learning for anomaly detection: A systematic review. *Ieee Access*, *9*, 78658–78700.
Nikolaev, A. S. (2019). Key approaches in determining role of the patent analytics in organization and stimulation of the innovation activity. *Innovation*, *12*(254), 119–123.
Özsungur, F., Tetik, N., & Kanat, E. (2023). Bankruptcy and Reorganization. *In the Digital Business Era. IGI Global*.
Park, H., Kim, K., Choi, S., & Yoon, J. (2013). A patent intelligence system for strategic technology planning. *Expert Systems with Applications*, *40*(7), 2373–2390.
Rahmani, A. M., Azhir, E., Ali, S., Mohammadi, M., Ahmed, O. H., Ghafour, M. Y., Ahmed, S. H., & Hosseinzadeh, M. (2021). Artificial intelligence approaches and mechanisms for big data analytics: a systematic study. *PeerJ Computer Science*, *7*, e488.
Rehman, S. U., Bresciani, S., Ashfaq, K., & Alam, G. M. (2022). Intellectual capital, knowledge management and competitive advantage: a resource orchestration perspective. *Journal of Knowledge Management*, *26*(7), 1705–1731.
Ryabokon, M. S., Skuybin, B. G., & Shcheglov, D. K. (2019). Patent Research as a Tool for Analysis of the Technical Solutions Market. *Administrative Consulting*, *11*.
Sekerin, V. D., Slepov, V. A., Gayduk, V. I., Bank, S. V., & Kravets, E. V. (2021). Blockchain technology development as tool for enhancing security in management and protection of intellectual property rights in additive manufacturing. *Revista Geintec-Gestao Inovacao E Tecnologias*, *11*(2), 1184–1200.

Shang, H., Liu, C., & Wang, R. (2022). Measurement methods of 3D shape of large-scale complex surfaces based on computer vision: A review. *Measurement*, *197*, 111302.
Shao, B., & Wang, H. (2025). Digital Economy, Industrial Structure Advancement and Human Capital Accumulation. *Finance Research Letters*, 107727.
Shuai, J., Peng, X., Zhao, Y., Wang, Y., Xu, W., Cheng, J., Lu, Y., & Wang, J. (2022). A dynamic evaluation on the international competitiveness of China's rare earth products: An industrial chain and tech-innovation perspective. *Resources Policy*, *75*, 102444.
Szczepańska-Woszczyna, K., & Gatnar, S. (2022). Key competences of research and development project managers in high technology sector. *Forum Scientiae Oeconomia*, *10*(3), 107–130.
Taherdoost, H. (2023). Deep learning and neural networks: Decision-making implications. *Symmetry*, *15*(9), 1723.
Tallarico, S., Pellegrini, L., Lazzarotti, V., & Lazzini, S. (2025). Boosting firms' absorptive capacity: the digital technologies edge. *European Journal of Innovation Management*, *28*(6), 2558–2580.
Trippe, A. (2015). Guidelines for preparing patent landscape reports. *Patent Landscape Reports. Geneva: WIPO*, 2015.
Trippe, A. J. (2003). Patinformatics: Tasks to tools. *World Patent Information*, *25*(3), 211–221.
Wang, C.-C., & Chien, C.-H. (2025). Machine Learning for Industrial Optimization and Predictive Control: A Patent-Based Perspective with a Focus on Taiwan's High-Tech Manufacturing. *Processes*, *13*(7), 2256.
Wang, J., & Hsu, C.-C. (2021). A topic-based patent analytics approach for exploring technological trends in smart manufacturing. *Journal of Manufacturing Technology Management*, *32*(1), 110–135.
Wen, X., Wang, J., Zhang, G., & Niu, L. (2021). Three-dimensional morphology and size measurement of high-temperature metal components based on machine vision technology: a review. *Sensors*, *21*(14), 4680.
WIPO. (2019). World Intellectual Property Indicators 2019. *Geneva: World Intellectual Property Organization*.
Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., & Qiu, C.-W. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, *2*(4).
Yang, D., Lee, J., Song, N. C., Lee, S., Kim, S., Lee, S., & Choi, S. (2023). Patent analysis on green hydrogen technology for future promising technologies. *International Journal of Hydrogen Energy*, *48*(83), 32241–32260.
Yavuz, M. S., & Çalik, H. (2025). AI and ML patent intensity and firm performance: A machine learning-based lagged analysis. *European Research on Management and Business Economics*, *31*(3), 100291.
Yun, S., Cho, W., Kim, C., & Lee, S. (2022). Technological trend mining: identifying new technology opportunities using patent semantic analysis. *Information Processing & Management*, *59*(4), 102993.
Zhao, A., & Wang, P. (2025). Predictive Framework for Regional Patent Output Using Digital Economic Indicators: A Stacked Machine Learning and Geospatial Ensemble to Address R&D Disparities. *Analytics*, *4*(3), 18.