Research on the Method of Integrating Knowledge Graph and Deep Neural Network for Automotive Technology Foresight

Lian Liu [1a], jun Wang [2b], dan Wu [3c], xingde Huang [3d], DeXiang Jia [2e], Aiqiang Pan [2f]

[a*]Corresponding author's e-mail: 872616590@qq.com , [b]kentywang@126.com, [d]huangxingde@sh.sgcc.com.cn, [c]wudan@sh.sgcc.com.cn , [e]jiadexiang71@163.com , [f]dsy_panaq@126.com ,

[1]State Grid Shanghai Electric Power Research Institute, Shanghai, 200437,China
[2]State Grid Shanghai electric power company, Shanghai 200120,China
[3]State Grid Energy Research Institute Co., Ltd., Beijing, 102209,China

Abstract. Realizing precise foresight of emerging technologies in the field of electric vehicles helps the company to advance its layout in this field, seize the technological commanding heights, and empower high-level technological self-reliance and self-reliance. This study is based on knowledge graph technology and constructs a technical knowledge graph based on the relationships and attributes between scientific and technological papers in the field. It revolves around the three main characteristics of novelty, social impact, and fundamental innovation in the field of electric vehicle technology, and constructs a complete and quantifiable indicator system for new electric vehicle technologies. The knowledge graph is used to extract feature values of each indicator, relying on deep neural network algorithms, Train the industry's emerging technology foresight model to achieve precise foresight for the development of new electric vehicle technologies. This study can provide valuable reference for technology foresight in the field of electric vehicles and support industrial development decisions.

Keywords: technological foresight; Knowledge graph; Electric vehicles, deep neural networks

1. Introduction

Technological innovation is a powerful engine for achieving high-quality economic development and the focus of leading enterprise development. In response, the 14th Five Year Plan for National Economic and Social Development of the People's Republic of China and the Outline of Long Range Objectives for 2035 point out that "we should focus on seizing the opportunities for future industrial development, cultivating leading and pillar industries, and promoting the integration, clustering, and ecological development of strategic emerging industries. New technologies further demonstrate the "head goose effect" in the process of industrial innovation and development, giving rise to an increasingly important space for dominance in this field, becoming the key to determining the company's future and competitive advantage[1]. Therefore, how to grasp the overall trend and seize the opportunity in the process of technological innovation, and accurately grasp the technological development trend of the

ICEKIM 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2337248

electric vehicle industry in the field of electric vehicles, is an important issue in the scientific and technological strategic layout of power companies.

Currently, scholars from various fields both domestically and internationally have conducted numerous studies on the issue of foresight for emerging technologies [2,3]. In the initial research stage of emerging technology foresight, the Delphi method method [4] and the analytic hierarchy process [5] are often used, and valuable expert experience is used to propose indicators to describe the characteristics of emerging technologies, which provides a solid foundation for subsequent research. The prediction of emerging technologies based on measurement models is more based on the quantitative indicator system of past or current technological characteristics to achieve judgment and prediction. For example, reference [6] constructs a model for identifying emerging technologies based on attribute sets and attribute measurement theory, including two primary indicators for technology and market, and six secondary indicators; Reference [7] identified four indicators in the identification system: novelty, persistence, community, and growth. The emerging technology foresight based on literature metrology mostly uses the direct bow network, co occurrence and coupling network of scientific and technological papers or related scientific and technological papers for cluster analysis [8, 9], such as building the similarity matrix of scientific and technological papers, evaluating the prediction effect according to the cluster density and similarity, fitting and analyzing the bow curve of the time series bow data of the emerging technology theme. Overall, current research on the foresight of emerging technologies mainly focuses on macro and meso level control, lacking micro, refined, and comprehensive research. It focuses more on and describes the external characteristics and network structure characteristics of technology, and lacks research on the essential characteristics of technology; Due to the lack of correlation research on relevant resources, the interpretability of predicting emerging technologies is not clear enough.

Therefore, this paper fully considers the essential characteristics of technology, enhances the interpretability of emerging technology foresight, builds an emerging technology indicator system from the dimensions of novelty, application scope and development ability by relying on the Semantic information of the knowledge map network of scientific papers, and uses the deep neural network algorithm with strong feature extraction ability to form low-level features into more abstract high-level representation features, and then trains and forms an industry emerging technology foresight model, Realizing the foresight of emerging technologies in the industry. Through the connection between nodes and edges, the knowledge map of scientific and technological papers adds a variety of semantic relationships to the papers, which not only fully utilizes the Semantic information of the technical papers' network to deeply mine technical features, but also strengthens resource relevance to improve the predictability and interpretability of emerging technologies in the field of electric vehicles.

2. Framework of Knowledge Graph for Scientific and Technological Papers

The specific research logic of this article is shown in Figure 1, which defines the concepts, relationships, and attributes of industrial technology papers based on the scientific and technological paper dataset, and constructs a knowledge graph of industrial technology papers; Based on the indicator system of emerging technologies in the industry, map graph query statements to extract characteristic values of each indicator, and normalize all indicators in scientific and technological papers; Evaluate emerging technologies in the industry using development potential classification labels, and ultimately use a normalized prediction dataset to achieve industry emerging technology foresight and visualize it.

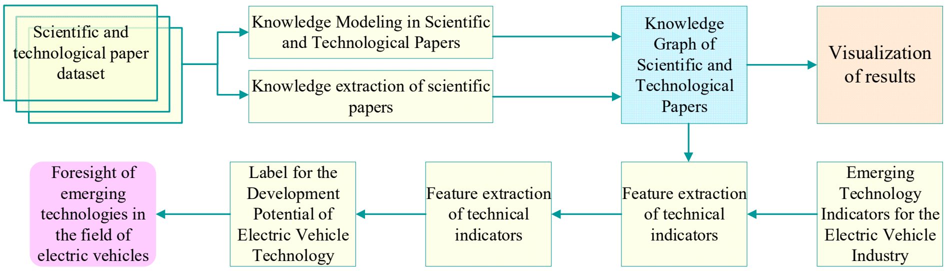

Figure1. Framework of Knowledge Graph for Scientific and Technological Papers

3. Construction of a knowledge graph for scientific and technological papers

Scientific and technological papers are an important carrier for technological innovation, achievement transmission and exchange, and have a significant impact on the development of related industries and technological breakthroughs. However, the data of scientific and technological papers are characterized by large scale and fragmentation, which brings great challenges to people's efficient and comprehensive understanding of data connotation. Therefore, this paper uses the knowledge map technology to realize the knowledge interconnection of scientific and technological papers [10], build the knowledge map of scientific and technological papers in the field of electric vehicles, and form a complex semantic network from the original fragmentation paper data through fusion and association to meet the perspective of human cognitive needs, so as to understand and analyze the integrity and relevance of scientific and technological paper data.

Firstly, obtain data from industry related scientific and technological papers through the China National Knowledge Infrastructure database, and analyze the semantic labels of technical paper data;

Then, using citespace software tools, the knowledge map of scientific and technological papers is modeled ontology, and the concepts, relationships and attributes of the concepts in the paper field are defined. The Resource Description Framework model is used to realize the standardized expression of the ontology in the corresponding technical field of the paper;

Then, according to the above designed scientific paper ontology, knowledge extraction is carried out on the obtained scientific paper data to obtain the entities and relationships of the instantiation of the scientific paper atlas;

Finally, the instantiated scientific and technological paper entities and relationships are shown in the form of attribute maps to form a complex graph structure semantic network, which can realize interactive query and association analysis, and provide research support for the subsequent construction of industrial emerging technology indicator system and industrial emerging technology foresight.

3.1. Knowledge Modeling in Scientific and Technological Papers

Ontology is a formal expression of concepts and their relationships, a standardized and formalized specification of shared conceptual models, and can endow information resources with more complete and clear semantics. Among them, domain ontology is a formal description of concepts and their relationships in important domains with limited scope. Based on the obtained industry related scientific and technological paper data, the main research object of this article is the concept of industry related scientific and technological papers and their relationships. From a model perspective, the domain ontology of scientific and technological papers includes classes, attributes, and relationships. After setting the attributes, the ontology model is visualized and displayed.

3.2. knowledge extraction of scientific papers

Entities and relationships are the basic units that constitute the knowledge graph of scientific papers. The entity of a scientific paper is an instantiation of the concept of the field of the paper, while the relationship is a description of the objective relationship between objects in the field of the scientific paper, defined as a certain connection between two or more entities. The knowledge of the knowledge map of scientific papers in this paper comes from the data of industrial related scientific papers obtained from the CNKI database. The exported data is saved in the ". xl-sx" format file. Each line of data record in the file contains multiple entity names and entity attribute information. According to the above knowledge modeling design of scientific papers, structured data that can be understood and calculated by computers are extracted from these data through knowledge extraction technology, Extract entities of different categories and map them to generate corresponding semantic relationships, thus achieving relationship extraction in the graph of scientific papers, as shown in Figure2. The entity attributes extracted in this article include: abstract, national economic category, author, title, publication time, DOI, keywords, etc.

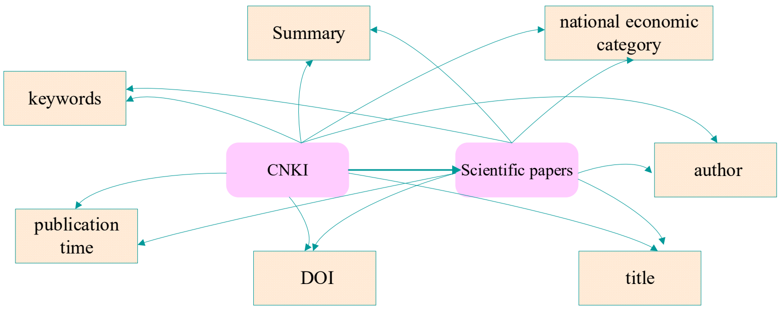

Figure2. The Entity Attribute Relationship of Knowledge Graph in Scientific and Technological Papers

4. Emerging Technology Foresight Model Based on Deep Neural Networks

By using the above method to annotate the classification data of the development potential of the paper, a training sample set of emerging technology foresight models in the field of electric vehicles is obtained. Furthermore, a relationship model between the indicators of scientific and technological papers and their development potential is constructed, namely the industry emerging technology foresight model, which is defined as:

$$y = f(x) \tag{1}$$

In which, y represents the development potential of industry related papers; X is the classification label system for the development potential of the paper; F is the industry emerging technology foresight model.

4.1. Deep neural networks Methods

Deep neural networks (DNNs) are neural networks with multiple hidden layers, consisting of several interconnected neurons. The layers of deep neural networks are fully connected, with the first layer being the input layer, the last layer being the output layer, and the middle layer being the hidden layer. Deep neural networks have strong feature extraction ability, can independently extract and learn target features, and have strong robustness and generalization ability, making them one of the widely used mainstream algorithms. The deep neural network model constructed in this article is shown in Figure 3. The input layer is based on the index features extracted from the knowledge map of scientific and technological papers, the hidden layer uses the cross entropy loss function and sigmoid activation function to improve the convergence speed of the DNN algorithm, and the predicted results of the output layer are compared with the actual result tags, so as to realize the application of the deep neural network algorithm in the prediction model of emerging industrial technologies.

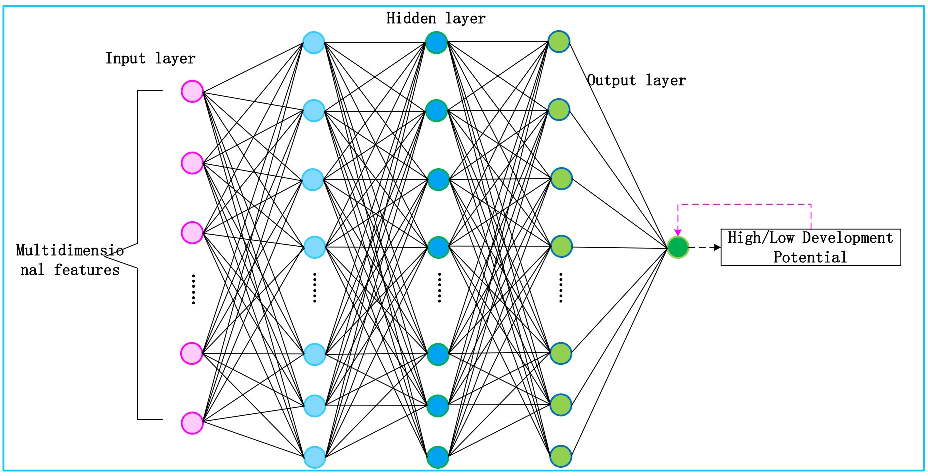

Figure3. DNN Model

4.2. Evaluation of Emerging Technology Foresight Model

The confusion matrix is used to record the forecast of emerging technologies in the industry, as shown in Table 1. The four main elements of the confusion matrix are used to characterize the prediction of industrial scientific and technological papers in the test set. Among them, true positive (TP) represents the number of correctly predicted industrial emerging scientific and technological papers that belong to industrial emerging technologies, and true negative (TN) represents the number of correctly predicted non industrial emerging scientific and technological papers that belong to non industrial emerging technologies, False positive (FP) indicates the number of papers that belong to non industrial emerging technologies that were mistakenly predicted as industrial emerging technologies, while false negative (FN) indicates the number of papers that belong to industrial emerging technologies that were mistakenly predicted as non industrial emerging technologies.

Table1. Emerging technology foresight confusion matrix

Anticipating Results of emerging technologies in the industry		**Anticipating Results**	
		High development potential	low development potential
Actual results	high development potential	true positive (TP)	pseudo negative (FN)
	low development potential	Pseudo Positive (FP)	True Negative (TN)

The evaluation was conducted using three indicators: precision (PR), recall (RE), and F-score. The indicator formula is:

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$RE = \frac{TP}{TP + FN} \tag{3}$$

$$PR = \frac{TP}{TP + FP} \tag{4}$$

$$F1 - score = \frac{2 \times PR \times RE}{PR + RE} \tag{5}$$

4.3. Technical foresight model training

Due to the wide variety and confusion of emerging technologies in the field of electric vehicles, as well as the large number of indicators, the training sample set of the industry emerging technology foresight model is constructed based on the feature extraction of indicators and their development potential classification labels based on the knowledge graph of scientific papers, and the training set and test set are divided. The paper data from 2015 to 2023 is divided into a training set and a testing set at a ratio of 3:1. The model training sample set is illustrated in Figure 4 The samples from the training and testing sets have an 11 dimensional vector:

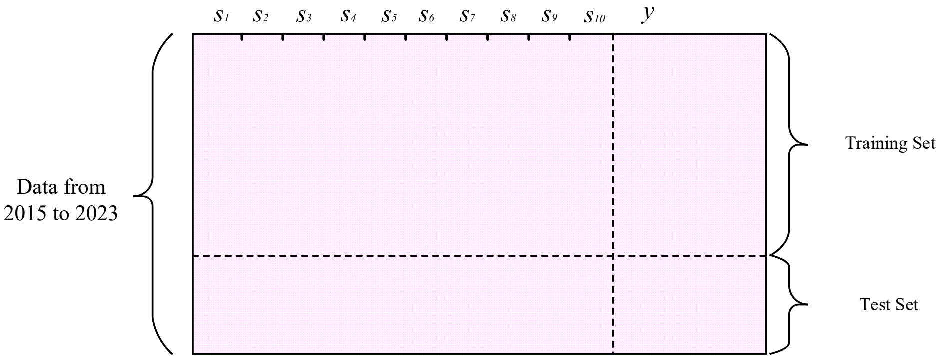

Figure4. Model training sample set

5. Current Status of Electric Vehicle Research

5.1. Published Statistics

Figure5 shows the publication situation and trends of articles on the topic of “electric vehicles” in China, which are included in China National Knowledge Infrastructure (CNKI) from 2014 to 2022. As can be seen from Figure5, research in the field of electric vehicles in China has shown a clear upward trend, with a total of 1,243 articles. There are 188 articles in the literature. With the follow-up development of China in the field of comprehensive energy, the research on related topics will become the focus of future research. It is expected that the number of articles on related topics will increase significantly in the future.

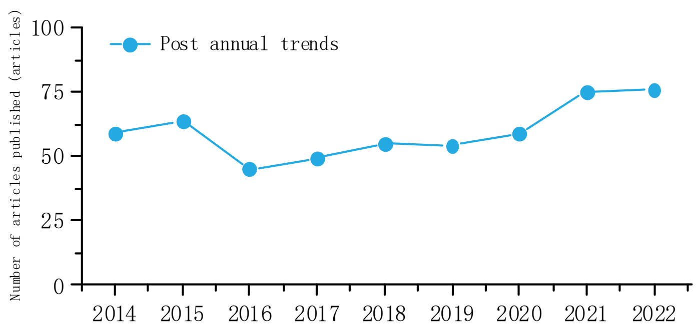

Figure5. Statistics published in the field of electric vehicles

5.2. Keyword co-occurrence analysis

Keyword This paper extracts the keywords of China's "electric vehicle" theme literature from 2014 to 2022, and removes some keywords that are not related to new technologies in terminal energy consumption. A total of 99 nodes are obtained and drawn as shown in Figure6. A knowledge map showing the co-occurrence of research keywords in the field of "electric vehicles" in China. The larger the circle and the more connections, the higher the research interest of the keyword in the discipline. Figure6 shows that in recent years, relevant research in the field of "electric vehicles" in China has mainly focused on themes such as ancillary services, energy storage, demand response, and load forecasting. Combining with Table 2, it can be seen that, except for the keyword "electric vehicle", the centrality of keywords such as "demand response", "virtual power plant" and "smart contract" all exceed 0.9, and the centrality of other keywords is between 0.7 and 0.9. It is also the core keyword in the research field of "electric vehicle".

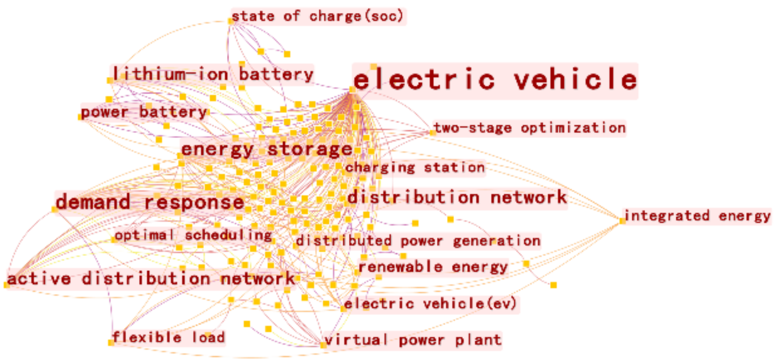

Figure6. Knowledge graph in the field of electric vehicles

Table2. Main keywords in the field of electric vehicles

Keywords	Centrality	Time/year
Demand Response	0.979	2020
Load Forecasting	0.897	2020
Virtual Power Plant	0.947	2020
Energy Management	0.88	2019
Smart Contract	0.919	2019
Battery Safety	0.96	2019
Master-Slave Game	0.801	2020
Orderly Charging	0.937	2020
Wireless Charging	0.982	2020
Air Conditioning Load	0.998	2022

5.3. Keyword Cluster Analysis

In order to fully sort out the research topics in the field of "integrated energy", this paper conducts cluster analysis on the keywords in the literature, and obtains the keyword clustering knowledge map as shown in Figure7,which mainly divided into 5 cluster centers, and the network homogeneity results are obtained to show its rationality. Combining Figure7 and Table2, it can be seen that the field of "Integrated Energy" is clustered in topics such as "Information-Physical Fusion System", "Carbon Neutrality", "Data Driven" and "Cogeneration", which cover a wide range.

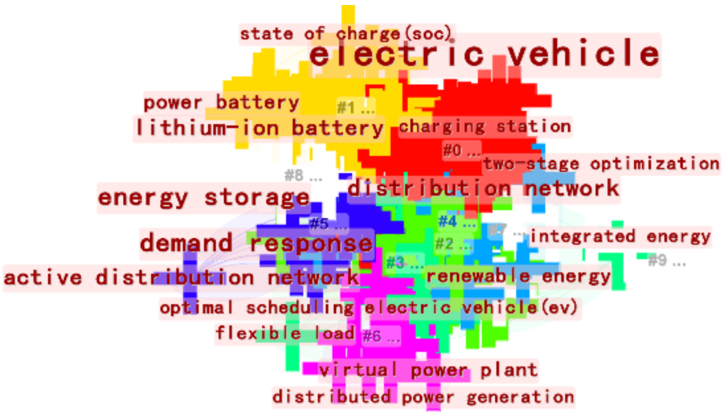

Figure7. Keyword cluster analysis

According to the main keywords included in each cluster, the research clusters of "distribution network", "load forecasting" and "battery safety" mainly focus on the impact of electric vehicle charging power on the operation of the distribution network and the complexity of charging load changes. The large increase in the number of electric vehicles has a very obvious effect on the operation mode of the distribution network, but at the same time, the large-scale use of electric vehicles also has a promoting effect on the reduction of oil vehicles, which makes the terminal energy consumption of this type increase sharply. It has an obvious positive effect on building a low-carbon society and achieving the dual-carbon goal, but the differences in different stages lead to fluctuations in the economy and low-carbon operation. Therefore, research in the field of "electric vehicles" provides the incubation of new technologies for terminal energy consumption. an effective way.

5.4. Knowledge Graph Indicator Feature Extraction and Technology Foresight Model Evaluation

Feature extraction is an important foundation for model training. Based on the constructed knowledge graph of scientific papers, indicator features are extracted and corresponding definitions are used to calculate the indicator scores of the dataset. Among them, in the development potential classification label, based on previous research experience, due to the large numerical span of the training dataset on various indicators, all paper indicators are normalized.

Build a DNN model for predicting emerging technologies in the electric vehicle industry. The DNN model obtained through multiple rounds of experiments has four hidden layers, including 10, 20, 20, and 10 neural units, and has two target classifications, namely high development potential and low development potential papers. The input of the DNN model input layer has the index characteristics of a 10 dimensional vector. The hidden layer will be initialized to a random value that conforms to the normal distribution. The predicted results of each training sample in the output layer will be compared with the actual development potential value. Through the continuous iteration of the loss function, the parameters will be optimized. When the loss function converges to a certain value, the model parameters will be optimized and the training will stop. From Table 3, it can be seen that the DNN model has good representational ability and overall performance in predicting emerging technologies in electric vehicles. During the training process, the DNN model can dilute irrelevant factors and better fit the complex nonlinear relationship between the indicator system and the development potential of new energy vehicle emerging technologies. It can not only comprehensively and accurately identify potential emerging technologies in the future new energy vehicle industry, but also minimize the omission of important emerging technologies in the industry, In order to avoid missing important "overtaking" opportunities in the process of industrial development, and provide decision-makers with more accurate and effective foresight of emerging technologies in the industry.

Table3. DNN Model Calculation Results

Accuracy rate (%)	Recall rate (%)	F-score (%)
98.1	94.6	93.3

6. Conclusion

This article relies on knowledge graph technology and focuses on research on predictive modeling of emerging technologies in the field of electric vehicles. It focuses on defining the technical concepts, relationships, and attributes of industrial science and technology papers, and constructing a knowledge graph of industrial science and technology papers. The network Semantic information is used to map the characteristic values of each index in the atlas according to the emerging technology index system of the industry, forming the emerging technology forecast in the electric vehicle field, providing valuable reference for the emerging technology forecast in this field, and providing decision-making support for the industrial development.

The main contributions of this paper are as follows: Extracting knowledge elements from the data source of scientific papers, realizing the standardized expression of domain ontology with the Resource Description Framework, and constructing the knowledge map of scientific papers according to the corresponding rule mapping. It is of great significance and reference value to further predict the emerging technologies of electric vehicles based on the knowledge graph of scientific and technological papers Based on the characteristics of emerging technologies in electric vehicles, construct a comprehensive, comprehensive, and accurate evaluation and analysis of scientific papers, and deeply explore the logical relationships based on the knowledge graph of scientific papers, laying a solid foundation for predicting emerging technologies in the industry.

The research results prove that the technology foresight formed by the fusion of knowledge graph method in this article has relative superiority, but there is still some room for deepening and expanding. The key issues of subsequent research are as follows: ① Only using scientific and technological paper data for emerging technology prediction, the data sources and types are not comprehensive enough, which may have a certain impact and deviation on the results. Therefore, the next step of research will be integrated into industry reports, relevant policies Market public opinion related data, realizing the identification and prediction of emerging industrial technologies driven by multi-source data, making the prediction results more comprehensive and representative; ② The knowledge graph of scientific and technological papers contains a large amount of exploitable knowledge and content. This article only extracts indicator features from the knowledge graph of scientific and technological papers. In the future, multi-dimensional, multi perspective, more in-depth, and fine-grained mining analysis will be conducted on the basis of the knowledge graph of scientific and technological papers, providing decision support for industrial development.

Acknowledgments: Supported by science and technology project of State Grid Corporation of China (Research on Key Technologies for Unified Application and Promotion of New Technologies in Terminal Energy Consumption 5400-202117406A-0-0-00).

References

[1] Belton, I., Cuhls, K. & Wright, G. (2022) A critical evaluation of 42, large-scale, science and technology foresight Delphi surveys. Futures & foresight science, 4(2),122-134.

[2] Gao, Z., Gao, M., Chen, C., Zhou, Y., Zhan, Z. & Ren, Y. (2023) Knowledge graph of wastewater-based epidemiology development: A data-driven analysis based on research topics and trends. Environmental science and pollution research international, 30(11), 28373-28382.

[3] Kim, H., Han, J. & Kim, Y. (2013) Study on Future Foresight of the Technology Commercialization Policy. Journal of Industrial Economics and Business, 26(2).

[4] Shen, Z., Ji, W., Yu, S., Cheng, G., Yuan, Q., Han, Z., Liu, H. & Yang, T. (2023) Mapping the knowledge of traffic collision Reconstruction: A scientometric analysis in CiteSpace, VOSviewer, and SciMAT. Sci Justice, 63(1), 19-37.

[5] Huang Lucheng&Lu Wenguang (2009) Research on Emerging Technology Recognition Based on Attribute Comprehensive Evaluation System Research Management, 30 (04), 190-194

[6] Liang Shuai&Zhao Lixin (2021) The connotation, challenges and implementation framework of technology foresight in the context of a risk society Science and Management of Science and Technology, 42 (03), 64-75

[7] Song Xinna, Guo Ying, and Xi Xiaowen (2020) Research on Multiple Indicator Emerging Technology Identification Based on Patent Literature Intelligence Journal, 39 (06), 76-81+88

[8] Yang Jie&Chen Kaihua (2021) Research on Priority Technology Selection for Social Vision and Challenges Scientific Research, 39 (04), 673-682

[9] Zhang Shuo, Wang Xuefeng, Qiao Yali&Liu Yuqin (2022) Current Status, Trends, and Future Reflections on Technology Prediction Research: A Data Analysis Perspective Library and Information Work, 66 (10), 4-18

[10] The 14th Five Year Plan for National Economic and Social Development of the People's Republic of China (2023) and the Outline of the Long Range Goals for 2035 by the Central People's Government of the People's Republic of China