Scientometrics (2024) 129:1975–2005
https://doi.org/10.1007/s11192-024-04961-0

Mapping and comparing the technology evolution paths of scientific papers and patents: an integrated approach for forecasting technology trends

Peng Liu[1] · Wei Zhou[1] · Lijie Feng[1,2] · Jinfeng Wang[2] · Kuo-Yi Lin[3] · Xuan Wu[1] · Dingtang Zhang[4]

Received: 30 April 2022 / Accepted: 5 February 2024 / Published online: 1 March 2024

Abstract
Exploring the key technology evolution paths in specific technological domains is essential to stimulate the technological innovation of enterprises. There have been many methods to identify the technology evolution path, but many of them still had some limitations. Firstly, many studies consider only a single type of data source without analyzing and comparing multiple data sources, which may lead to incomplete evolution paths. Secondly, the text mining methods ignore the semantic relationships between technical terms, making path tracing inaccurate. In this study, we develop an integrated approach for mapping the technology evolution paths of scientific papers and patents. To better forecast the technology development trends, the gap analysis between scientific papers and patents and the identification of potential topics are also applied. The all-solid-state lithium-ion battery technology is selected for the empirical study and the related technology evolution trends and the technology opportunities are focused on. The empirical case research results show the proposed method's validity and feasibility. This method can be helpful for understanding and analyzing the specific technology, which provides clues for forecasting technology development trends in enterprises. Furthermore, it contributes to the coordination of research and development efforts, which provides a reference for enterprises to identify technology innovation opportunities.

Keywords Technology evolution path · Technology opportunities analysis · Technology trend · SAO (subject-action-object) semantic analysis · Topic modeling

✉ Jinfeng Wang
wangjinfeng@shmtu.edu.cn

1 School of Management, Zhengzhou University, Zhengzhou 450001, China

2 China Institute of FTZ Supply Chain, Shanghai Maritime University, Shanghai 201306, China

3 School of Business, Guilin University of Electronic Technology, Guilin 541004, China

4 Zhengzhou Coal Mining Machinery Group Co., Ltd, Zhengzhou 450001, China

Introduction

The rapid changes in science and technology shorten the process of technological innovation, which provides tremendous opportunities and competitive advantages for enterprises. The ability to excavate deeply and judge precisely future technological opportunities has become vital for enterprises (Ren & Zhao, 2021). How to detect and predict future technology trends given a better understanding of their emergence is also a key concern for both countries and enterprises.

As an important tool to track the development and evolution of technology, the technology evolution paths can improve the accuracy of technology opportunities identification. At present, the definition of technology evolution is not unified (Liu et al., 2022a, 2022b). Most researchers consider that the technology evolution refers mainly to the evolutionary changes in the specific technology over time. For the targeted technology area, the technology evolution generally encompasses all technological activities and the accumulation of innovations in the field from its inception to its current stage (Liu et al., 2020). Specifically, the technology evolution pathways not only reflect the nature of technological innovation but also effectively predict technology opportunities for problem-oriented (Huang et al., 2021). Therefore, clarifying the technology evolution paths is crucial to stimulate the technological innovation of enterprises through technology forecasting.

In the study of technological evolution path, discovery and linkage of technological topic information is a crucial problem. The diversified approaches have been proposed to identify technology evolution paths, including International Patent Classification (IPC) analysis, citation network analysis, and co-word analysis. Although the series of analysis methods for technology evolution paths are mature, there are still some limitations such as (1) The structured information such as patent codes and citation information are used as analysis data, ignoring the key technical terms contained in unstructured information, which makes it difficult to obtain fine-grained technology evolution results (Han et al., 2021; Lee et al., 2021). (2) The relationships between technical terms are usually focused on word frequency analysis rather than semantic analysis, which may not be sufficient to reveal the changes in the technology trends (Li et al., 2019).

To overcome these shortcomings, an integrated approach using SAO semantic analysis combined LDA topic modeling to process data in the targeted technological field is proposed. The SAO structure can effectively acquire the inner connections between independent words to obtain semantic linkages between technical terms. This approach makes up for the shortcomings of existing studies in the semantic analysis of technical terms, which provides complete connectivity information for the evolution paths.

As the two foremost forces jointly determining the development of technology innovation, science and technology contribute to the development of the technology innovation together. The relationship between science and technology is complex, and there is a one-way flow from basic scientific research to applied research and products (Xu et al., 2020). However, the non-linear relationship where technological progress drives scientific progress may also exist (Ba & Liang, 2021). Since the different purposes, statement, and quality of scientific papers and patents, there are certain commonalities and specialties between them (Xu et al., 2021). The analysis and comparison of scientific papers and patents can be tracked to discover technological opportunities and construct high-quality roadmaps (Shen et al., 2020). Currently, some studies adopt patents or scientific papers alone as the data source for technology opportunity analysis, whereas few studies combine these two sources for evolution analysis (Liu et al., 2022a, 2022b). To some extent, only analyzing scientific papers or patents is not enough to deeply understand the future trend of technology (Hajikhani & Suominen, 2022). Therefore, we explore and map the high-quality technology evolution paths based on scientific papers and patents. The final technology trends and opportunities are obtained based on the gap analysis and potential topic analysis between scientific papers and patents, which provides a more overall perspective into evolution research.

In summary, the key contributions of this paper can be summarized as follows:

1. In identifying technology evolution pathways, this paper proposes an integrated method combining LDA topic clustering model and SAO semantic analysis methods to mine specific technology terms and technology topics. Comparing with the traditional LDA topic clustering method, this method can clearly represent the technology topic information and its changes in different periods in the context of semantics, and can also effectively identify the "problems" and corresponding "solutions" in technical information. Finally, this paper takes the field of all-solid state lithium battery as an example to conduct an empirical study to verify the effectiveness of the proposed method. Meanwhile, it also points out the direction of future technological innovation in the field of all-solid state lithium battery.
2. This paper uses patent and scientific papers data sources and analyses the evolutionary pathways and clustering topics between different data sources. The gap analysis and potential topics analysis are also designed to predict the technology trends and technology opportunities in the future, which analyses the technology trend from the relevance and difference of multi-data sources.

The rest of this paper is organized as follows. The "Related work" section briefly reviews the related works. The "Methodology" section presents the overall research process and method introduction. The "Empirical study and results" section conducts a case study in the field of all-solid-state lithium-ion battery research. The "Discussion and Conclusions" section lays out our key findings and future works.

Related work

Technological opportunity analysis based on multiple data sources

Most of the previous studies on technological opportunity analysis revolve around patent data, which represents the contributions of technological development and innovation. While in the knowledge-intensive technologies, scientific research is the basis for technological development and is considered the origin of technology and innovation (Lee et al., 2015). As the main research method to disseminate and communicate the scientific knowledge, the scientific papers reflects the process of the scientific development and the results achieved La et al. (2020) It is thus clear that science and technology are inextricably linked. At present, there are also many scholars who conduct technology opportunity analysis from a science-driven perspective and apply the results of scientific research to technology development.

Previous scholars have conducted technological opportunity analysis by mixing two kinds of data or using two kinds of data separately. The related studies can be broadly classified into three categories.

The first type of study mainly focuses on the mixture use of scientific papers data sources and patents, which does not distinguish between them. For example, Viet and Kravets (2022) analyze the technological trends in asset performance management in the smart energy sector using appropriate data mining methods through extensive experiments on scientific papers and patent data. Balland and Boschma (2022) identifies regions at the frontiers of science and technology based on patents and scientific literature. Meanwhile, based on the identification results, an in-depth analysis of the question of whether scientific capabilities within a region can be translated into technological leadership is carried out. Shi et al. (2022) draw on scientific papers and patents related to onboard hydrogen storage systems (OHSS) and develop a novel methodology for investigating the past, present, and future development trends in OHSS. The second type of study mainly focuses on the identification of technological opportunities using dissertation data and patent data separately, and overall analysis of the combination the identification results. For example, Wang et al. (2015) applies text mining and a descending dimensional clustering algorithm to cluster papers and patents within the field of micro-algal biofuels. Finally the potential technological opportunities are explored in both theory and practice. Jiao et al. (2023) uses the RAKE algorithm to extract keywords from scientific papers and patents and calculates the cosine similarity between cluster technology domains. And finally the potential technology opportunities are identified by calculating the semantic similarity of the cluster technology domain vectors. Leon-Silva et al. (2020) Perform co-author network arrays, patent research and commercial tracking by analyzing silver nanoparticle-related scientific papers and patents using bibliometric methods. Finally, the technology transfer pathway of silver nanoparticles from academic part to industrial application is shown in the paper, which provides opportunities for research, development, application and academic collaboration in the field of nano-materials.

The third type of study mainly focuses on forming scientific and technological topics based on scientific papers and patents separately for comparative analysis. This type of study overcomes the shortcomings of the first two types of studies that do not take full advantage of the relationship between science and technology. For example, Li et al. (2023) proposes a new technology opportunity identification method that uses scientific papers and patents as data resources, and integrates SAO semantic mining and outlier detection method. The method can effectively and comprehensively identify technology opportunity from the two levels of technical problems and technical solutions. Feng et al. (2023) applies the LDA model to identify the topics between papers and patent data. All of topics are divided into several innovation dimensions according to their attributes and the TEMPEST model to identify specific technology opportunities of proton exchange membrane fuel cell. Takano and Kajikawa (2019) construct the scientific paper and patent citation networks separately and calculate the cosine similarity between the paper and patent clusters. Then, the emerging clusters with low similarity between them are selected for technology opportunity analysis. Hernandez-Quintanar and Rodriguez-Salvador (2019) apply the competitive technology intelligence methodology to analyze the scientific papers and patents in the 3D printing field. The main technology trends regarding materials and uses are determined according the analyzing results, which gives a new window of opportunity for exploring the use of 3D bio-printing in a new area.

To sum up, there is a close relationship between scientific research and technology development, and scientific research has a driving force for technology development. Comparing with the using a single scientific papers or patents, combining the scientific paper data and the technology patent data will facilitate a more comprehensive technology opportunity analysis. However, many studies have conducted joint analyses of papers and patents with separate analyses of the two data sources, despite the fact that they use both data and cover more scientific and technical information. This may result in a lower degree of correspondence between the clusters based on the two data sources and a coarser technical granularity in the final result, which is not conducive to mining more technological opportunities.

Technology evolution path analysis

Technology evolution refers to the evolutionary changes in a specific technology field over time, which can predict future trends in technology. For the specific technological domain, the technology evolution generally encompasses all technological activities and their accumulation of innovations from the initial to the current stage of the field (Liu et al., 2020). As the visualization of technology evolution results, technology evolution paths can not only map the nature of technological innovation but also effectively predict problem-oriented technology innovation opportunities (Coccia, 2019). Therefore, clarifying the technology evolution path is crucial to stimulate the technological innovation of enterprises.

Currently, numerous scholars have studied technology evolution paths using diversified methods, including patent classification codes-based, main path-based, and text mining analysis methods. Firstly, the research based on the International patent classification (IPC) analysis method mainly explores the technology evolution pathways through the technology classification among different IPC codes (van der Pol & Rameshkoumar, 2017). For instance, Fernandes et al. (2020) discussed the solid oxide fuel cell technology paths in the United States and Japan based on the analysis of relevant IPC codes. Secondly, the main path-based analysis method mainly tracks the knowledge trajectories and correlations in complex citation networks to find out the main technology evolution pathways (Lai et al., 2021). For instance, Kumar et al. (2020) focused on the mobile payment system innovation and identifies the main evolution process using the main path analysis. Kim et al. (2021) developed an improved knowledge persistence-based main path approach to identify technology evolution paths of block chain technology. Thirdly, the text mining methods mainly analyze unstructured data efficiently in large datasets to identify areas in the wider innovation research landscape (Chen & Ho, 2021). For instance, Li et al. (2020) designed a hierarchical visual analysis system and combined the advanced text mining technologies, which could help comprehend the evolution of one discipline. Miao et al. (2020) used a novel approach based on text-mining to identify technology evolution pathways in 3D printing technology field.

However, there are still some limitations in using the above methods for technology evolution path analysis. In terms of data information extraction, only using structured information (such as IPC, citation information) ignores the important elements contained in unstructured information (such as text abstract, semantic information), which makes it difficult to obtain more detailed technology opportunities. In terms of semantic analysis, text mining-based methods mostly focus on the single word rather than semantic relationships, which may lead to inaccurate linkage analysis. Therefore, in this study, we use SAO semantic analysis to process data information in the target technology domain (He et al., 2019). A more complete semantic understanding can be obtained by deeply analyzing the inner connections between independent phrases.

Subject-action-object (SAO) semantic analysis

SAO semantic analysis is developed based on TRIZ theory, which can effectively reflect the semantic relationship among different technology elements (Wang et al., 2017). Furthermore, SAO semantic analysis can form a problem–solution model for the identification of core components in the technology innovation process (Yang et al., 2017). In this model, the S structure represents the new solution, the A–O structure represents the old technology point or the problem to be solved (Kim & Yoon, 2021).Therefore, the SAO semantic analysis is not only able to show the semantic relationship between technical elements, but also to mine the actual technical question and corresponding solutions in a more complete way.

SAO structures also can be visualized to depict the technology evolution paths. This method effectively makes up for the shortcomings of the above methods, which facilitates the track of technological elements and the technology trends. Now, this method is widely used to identify technology opportunities and technology components and to calculate patent similarity. For instance, Kim et al. (2018) developed the new method SAO-x, which enabled an in-depth examination of the purpose and effect of the technology. Choi et al. (2011) presented a procedure that formulates an SAO network by using SAO models extracted from patent documents and demonstrates its effectiveness. Wang and Yang (2019) calculated the weight of the SAO in patents in the robotics area and measure the similarity between patents by introducing the calculation index DWSAO.

In this paper, a hybrid method to construct technology evolution maps is proposed by identifying specific SAO relationships. The SAO semantic analysis is used to determine the semantic association relationship between technical elements in the specific technological domain. Identifying meaningful links between two SAO structures results in more accurate obtainment for technology opportunities.

Research methodology

In this section, the whole research process is summarized. The process includes the use of the SAO semantic analysis based on LDA topic model to capture semantic relationships among topics, the construction of the technology evolution pathways based on scientific papers and patents separately. The gaps analysis and the potential technical topics analysis are also introduced to identify technology trends.

The framework of the integrated method for technology evolution is shown in Fig. 1, which includes four phases.

Data collection and preparation

The field of all-solid-state lithium batteries technology (ASSLIBs) is selected for a empirical study. In recent years, with the rapid development of the electric vehicle era, ASSLIBs are being considered as attractive and promising technologies for next-generation clean energy storage (Ke et al., 2020). Therefore, analyzing the evolution trends of ASSLIBs technology is of great importance for R&D personnel in chemical industry enterprises and policymakers.

The web of science (WOS) and Derwent innovations index (DII) databases have been selected as the data sources for collecting data. Based on literature surveying and

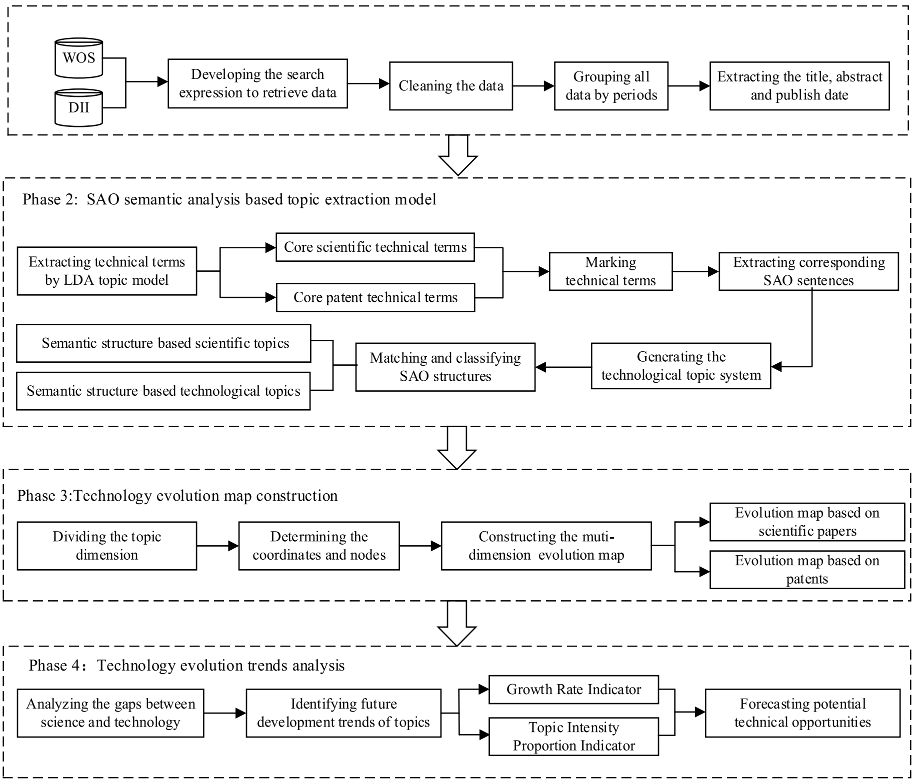

Fig. 1 The framework of mapping the technology evolution paths

expert knowledge, the search queries "((all-solid-state lithium batteries) or (all-solid-state lithium-ion batteries) or (all-solid-state Li-ion batteries))" is used to search the scientific papers from Web of science, and after manual elimination of scientific papers not related to all-solid-state lithium batteries, 3274 scientific papers are retrieved from 2002 to 2021 (the search time is January, 2022). The types of articles include Articles, Proceedings Paper and Reviews. The search term "((ASSLIBs) or ((all-solid-state) and (lithium-ion batteries)) or (all-solid-state Li-ion batteries)) is used to search the patents from Derwent Innovations Index database, and 1876 patents are retrieved from 2002 to 2021 (the search time is January, 2022).

During the data preprocessing, according to the overall time span based on the time interval proposed by Liu et al. (2020), the scientific papers and patents datasets are divided by year and the overall time scale is divided into five periods. Among them, each period includes four consecutive years. The final period setting and the corresponding number of papers are shown in Table 1. Second, the word-tokenize tool in NLTK natural language package of python program is used to is used for word segmentation. Then the POS tagging is also performed on the divided datasets. Meanwhile, unifying the singular and plural number, merging synonym, transforming full name and abbreviation, and removing irrelevant data are performed in turn. In specially, we remove some publisher information (such as "(C) 2018 Elsevier B.V. All rights reserved."), punctuation (such as "%"、 "("、 "<"), and other irrelevant characters (such as "this" "paper")

Table 1 Number of documents each period

Time period	T1	T2	T3	T4	T5
Year span	2002–2005	2006–2009	2010–2013	2014–2017	2018–2021
Number of scientific papers	110	130	275	517	2242
Number of patents	2	25	197	480	1172

from the data based on the constructed stop word list, which ensure objectivity and reasonability of the data.

As shown in Table 1, with the technology's development and social demands entering a phase of sharp increment, the number of scientific papers and patents grows rapidly each year. Particularly, the number of scientific papers and patents increased sharply in T5 period. (333.62% increase for papers and 144.17% increase for patents compared with T4 period). These growing rates show that the research and development of ASSLIBs have been energetic in last four years. Besides, the number of scientific papers has been more than the number of patents since 2002. This indicates the development process of acquiring scientific results has been faster than the process of applying the related technologies.

SAO semantic analysis based on LDA topic model

In this phase, the LDA topic model is firstly used to determine the number of topics and the corresponding technical terms of scientific papers and patents. Then the technical terms above are tagged in the raw datasets and the SAO semantic analysis method are used to extract the semantic structures corresponding to each technical term. Finally, through semantic similarity calculation, the SAO structures with high similarity are clustered to obtain the final technical topics and technical elements, which provide the data basis for the construction of the technology evolution pathways. This phase is conducted using three specific tasks which are applied one after the other. The three detailed tasks are as follows.

Extracting technical terms based on LDA topic model

Based on the acquired data under the time slice, the appropriate clustering number is firstly determined by calculating the function perplexity (Sugimoto et al., 2011). The perplexity is a popular indicator of the LDA model to evaluate the language model, applied to automatically evaluate the quality of topics. This paper is mainly based on the Gensim library in Python to calculate the perplexity value for different periods to ensure the optimal number of topics, where the smallest perplexity value indicates the best performance of the model. The mathematical formula for Perplexity calculation is as follows.

$$Perplexity(D) = \exp - \frac{\sum_{d=1}^{M} \log(p(w))}{\sum_{d=1}^{M} N_d} \tag{1}$$

where D denotes the whole data set, M denotes the number of texts contained in it, $\sum_{d=1}^{M} N_d$ denotes the number of words in the data, and the $p(w)$ denotes the probability of that document generating a word in the data.

Then set the α and β parameters of the Dirichlet which denote the per-document topic distribution and the per-topic word distribution. Finally, the technical topics under different periods are obtained after multiple topic training. The core of the LDA topic model is using the Dirichlet distribution to reduce the dimension of documents and achieve efficient clustering of documents and terms. The specific process is to first construct the document-topic matrix by generating topics of each document, and secondly construct the topic-word matrix by generating corresponding word distribution of each topic. The above process is repeated until the topic and subject words of all documents are generated. The core formula is as follows.

$$p(\theta w_i|\alpha, \beta) = p(\theta|\alpha)\coprod_{n-1}^{N} p(z_n|\alpha)p(w_n|z_n, \beta) \tag{2}$$

where d represents the current patent document, k represents the generated topics, which obey the Dirichlet prior distribution $\theta_{\mathrm{d}} \sim Dir(\alpha)$. $Zdi \sim Multinomial(\theta_d)$ represent the topic of word i in document d, $\phi_k \sim Dir(\beta)$ represents the distribution of lexical items generated for each topic k, and $w_{di} \sim Multinomial(\phi z_{di})$ represents the lexical item of word i in patent d. For a detailed explanation of the LDA, researchers can refer to Jelodar et al. (2018).

For this process, firstly the best appropriate term clustering number is obtained by calculating the perplexity of the scientific papers and patents datasets. Secondly, the LDA topic model is run to get the several technical terms clustering results based on scientific papers and patents. However, there are still some problems when using the LDA topic model results for analysis: simple noun words/phrases contain limited semantic information and may belong to several different technical topics. Therefore, without understanding the semantic context in which the technical terms are embedded, the researcher is unable to accurately interpret the technical topics at a certain depth and breadth, resulting in inaccurate topics recognition. Therefore, in this paper, we mark the technical terms obtained from the LDA topic model in the raw datasets, and then use the SAO semantic analysis method to extract corresponding SAO structure. This method improves the ambiguity of determining technical topics relying only on the LDA topic model, and improves the semantic relevance and accuracy of technical topic identification.

Extracting the SAO structure and topic clustering

Based on the identified technical terms, we use the Stanford Parser software to mine the SAO structure of the sentence in which the technical terms are located in the raw datasets. When the characteristics of SAO are applied to technical topic analysis, the relationships between technical terms are specified as the technology "problem" and corresponding "solution". This allows researchers to mine information not easily identified by common literature analysis methods such as bibliometrics and citation analysis, and provides researchers with more detailed and accurate technical topic analysis and trend prediction. In this paper, the specific SAO structure extraction process is roughly divided into two steps to meet the special requirements for natural language processing. The more detailed steps are shown in Table 2.

In this process: a. expand the technical terms lists based on morpheme, obtaining nouns, verbs, participles and other forms of technical terms to ensure the integrity of

Table 2 SAO structure extraction steps

Steps	Detailed explanation
Steps 1: technical terms expansion and tagging	
a. Expand the technical terms lists	Expand the topic-related terms/phrases lists based on lexical expansion methods (including, noun, verb and other forms) and the identification of core paragraphs
b. Mark the expanded technical terms	Complete the tagging of technical terms in the original datasets using python program to provide retrieval guidelines for the next step of SAO structure extraction
Steps 2: SAO structure extraction	
a. Extract the SAO structure	Extract the SAO structures based on Step 1 using Stanford Parser software for technical terms
b. Clean and consolidate	Clean and merge SAO structures using fuzzy matching and subject headings

SAO structure screening. b. the python program is used to tag sentences containing technical terms. c. SAO structures are extracted from tagged sentences containing technical terms based on the Stanford Parser tool. d. we clean the SAO structures based on expanded technical terms and stop word lists, and integrate of SAO structure based on fuzzy matching. It is worth noting that we should filter and retain the valid SAO structure. The extracted SAO structure should be directly related to the studied technical terminology, which corresponds to the S structure or O structure in the SAO structure. In this way, the SAO structures obtained finally can preserve their semantic features while emphasizing technical terms.

The Scientific papers and patents are usually used to solve problems in manufacturing or research and to provide corresponding solutions. A scientific paper or patent can be composed of multiple topics, each of which can be described by a series of "problems" and their "solutions" related to that topic. Compared with single word and phrase, the SAO structure takes into account the specific semantic information, i.e., the "problem" of a technology and the corresponding "solution", which facilitates topic identification and interpretation.

Therefore, in this paper, we construct the SAO structure matrix based on the extracted SAO structure and introduce it into the LDA topic model as input data. Then, the SAO structures are clustered again based on the LDA topic model operation principle and Eqs. (1), (2). Finally, we complete the technical topics based on the clustering results. The detailed LDA topic model based on the SAO structure are shown in Fig. 2.

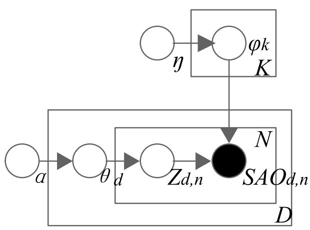

Fig. 2 The LDA topic model based on SAO

Generating the technological subsystems

After obtaining the technical topics and the corresponding SAO structures, the technology systems are constructed to provide data information for the construction of subsequent technology evolution pathways. The technical topics system based on the SAO structure is shown in Fig. 3.

The specific steps are as follows. First, based on the technical topics and the structure of the battery components, all SAO structures are divided into three categories: electrodes, electrolytes and interfacial layers. These three categories constitute the first level of the technology system, which can be defined as the content of the vertical axis. Second, the technical topics are used as the technology evolution topics, and the technology elements S or O in the SAO structure are used as the development node in this topic. And the association relationship among the technology elements is judged by the A structure in SAO sentences to highlight the inheritance and improvement relationship between the technology elements. These nodes from the S and O structure and the A structure serves as the inter-node linkage, which constitute the second layer of the technology subsystem.

Technology evolution map construction

The combination of the topic clustering method and SAO semantic analysis method is used to map the evolution path of science and technology topics. This hybrid method can not only clearly show the evolution relationship between terms, but also be more systematic and comprehensive than the single method analysis. In this paper, all technological topics and terms from SAO structures are divided into different dimensions for generating the subsystem of the whole ASSLIBs technology. In the evolution map, the vertical structure indicates the division of the technological subsystem, and the horizontal axis indicates successive slices of time. The detailed steps to construct the map are as follows.

Firstly, based on the subsystem of the ASSLIBs technologies, we classify the SAO structures into appropriate subsystem layers, and locate the topics and related terms on the map. Secondly, the topics or related terms with semantic relations are connected and the relationship between them is marked on the line. Finally, we construct the evolution pathway mappings of ASSLIBs technologies based on the scientific papers and patents separately.

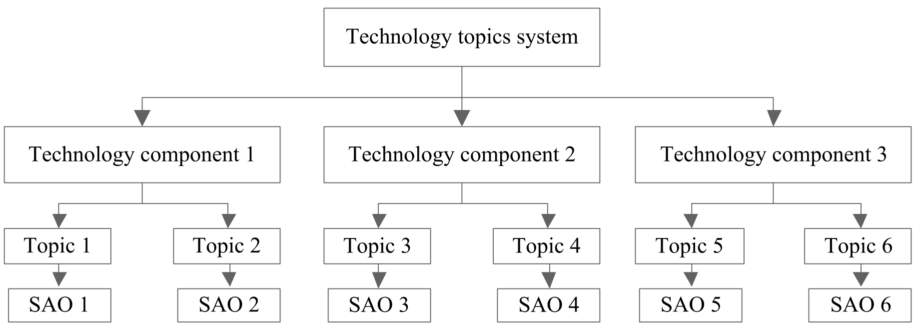

Fig. 3 The technical topics system

Technology evolution trends analysis

After the technology evolution pathway mappings based on the scientific papers and patents separately, we carry out the difference analysis of the technological topics in recent period, so that the potential technology opportunities could be forecasted. Secondly, the topic status is determined in the most recent phase to tap potential topics and served as a technical opportunity for future development. In this paper, the potential topics are determined by growth rate indicator and topic intensity proportion at this stage.

To be specific, the definitions of the two parameters are listed as follows:

(1) Growth Rate Indicator (GRI): This indicator reflects the growth rate of the topics. The topic life state can be judged by comparing the number of documents on one topic in this period with the sum of documents on this topic in five years. If the value increases over successive years, it indicates that the topic is in the newborn or growth stage. This paper refers to the study of Li et al., (2019), the GRI is calculated:

$$GRI_{\mathrm{i}} = \frac{a_{\mathrm{t}}^{i}}{\sum_{t-4}^{t} a_{t}^{i}}, (t \geq 4) \tag{3}$$

where a_t^i represents the document number of topic i in period T.

(2) Topic Intensity Proportion Indicator (TIP): This indicator reflects the topic popularity and heat, which can be used to reveal the development potential of a topic in the future. Specially, in the $t+1$ period, the heat of a certain topic exceeds the average heat of different topics in the same period, and the intensity of the topic is higher than the average intensity of the topic. After the $t+1$ period, this topic can be considered as a topic with development potential and needs to be focused on. According to the (Li et al., 2019), first, the Average Topic Intensity is calculated with the following equation:

$$ATI_{\mathrm{t}} = \frac{\sum_{1}^{N} TI_{t}^{z}}{N} \tag{4}$$

In Eq. (4), the TI_{t}^{z} represents the intensity of topic Z in period T, which can be calculated by the topic probability model. N represents the number of research topics detected in period T. ATI_{t} represents the average topic intensity of all topics in the analysis data source within time period T. If $TI_{\mathrm{t}}^{z} > ATI_{\mathrm{t}}$, the topic is higher than the average topic intensity value and has an excellent development potential in the next stage. In order to better represent the evolution and development of the topic at different stages, the normalized processing method is selected and the intensity ratio index of the topic is introduced in Eq. (5).

$$TIP_{\mathrm{t}} = \frac{TI_{t}^{z}}{ATI_{t}} \tag{5}$$

In Eq. (5), Topic Intensity Proportion (TIPt) represents the normalization result of the topic intensity value and the average topic intensity value. If TIPt > 1, it indicates that the topic is higher than the average topic level and has better development potential in the future.

Results and discussions

SAO semantic analysis based on topic extraction

The topic extraction based on SAO semantic analysis consists of three parts: technical terms extracting based on LDA topic model, SAO structures extracting and analyzing, the topic clustering and the construction of the technological topic system.

Technical terms extraction based on LDA topic model

The corresponding titles and abstracts are extracted from each period data of the scientific papers and patents. The reprocessed data are converted into CSV format as input datasets in LDA topic modeling. First, the perplexity value of the model under different topic number is calculated, and the optimal number of topics in each period is obtained. Then, set the optimal number of topics in the LDA topic model. According to the Gibbs sampling method, the experimental parameters in the model are specifically set as follows: alpha = 0.01, beta = 0.01, minimum_probability = 0.001, chunksize = 100 and the max_iterations = 10,000. The specific data results are shown in Table 3.

Then, the LDA topic model is used to cluster the scientific papers datasets and patent datasets separately. The results of technical terms clustering are shown in Tables 3 and 4. In Tables 4 and 5, it reflects the technical terms and their probabilities under a technical topic under different periods.

Table 5 shows the patent technological terms with their probabilities for each period from T2 to T5.

At different periods, each topic contains 7–10 technical terms with maximum probability. The above technical terms are indexed in raw data by the specific python program.

SAO structure extraction and topic extraction

The top related technological terms from Tables 3 and 4 above are indexed in the original data sources by python program. The indexing process is shown in Table 2. Then we use the Stanford Parser software to identify SAO structures from the flagged data. Meanwhile we clean and merge SAO structures using fuzzy matching and subject headings. As a result, the SAO structures from scientific papers and patents data sources are extracted separately, and we obtain 378 and 289 groups of SAO structures. After extracting and cleaning the SAO structures, we use this processed data as an input into the LDA topic model and run it again to identify the technical topics.

In this process, we calculated the number of the similar SAO structures to distinguish and identify key technical elements. In a SAO structure, the verb part "Action" in SAO structure is as the relationship among topics which include *"increase", "improve", "reduce"* and so on. We invited four researchers from Lomon Billions Group Co., Ltd in Henan, China to screen and technological topics and the SAO structure. These four researchers have been researching in the ASSLIBs field for over 5 years and have rich theoretical and practical experience to help identify the technology topics with the help of LDA topic model. Synthesizing the expert's opinion, we identify the technical elements with high frequencies and core evolutionary relationships among the scientific topics and

Table 3 Perplexity calculation results in different time domains

Time period	Topic interval	The perplexity minimum	Corresponding optimal number of topics in scientific papers	The perplexity minimum	Corresponding optimal number of topics in patents
2002–2005 (T1)	[1–10]	36.753	2	–	–
2006–2009 (T2)	[1–10]	66.833	5	4.042	2
2010–2013 (T3)	[10–30]	145.015	9	125.096	6
2014–2017 (T4)	[10–30]	393.030	11	197.025	10
2018–2021 (T5)	[10–30]	529.820	15	241.860	14

Table 4 Scientific papers clustering results in ASSLIBs (partial)

Period	Top related technological terms	Probability	Period	Top related technological terms	Probability
T1	Polymer	0.0572788	T1	Thin-film	0.0700364
	Film	0.0570704		Cells	0.0663235
	Glass–ceramic	0.0509425		Film	0.0656183
	Polyethylene oxide	0.0472194		Polymer	0.0635012
	High-capacity	0.0472085		Rechargeable	0.0560169
	Electrode	0.0444099		Solid-state	0.0507405
	Solid polymer electrolyte	0.0424598		Al2O3-coated	0.0504098
	Sulfur	0.0424587		LiCoO2	0.0493432
	Li	0.0406044		Charge/discharge	0.0477206
				Cycles	0.0467331
T2	LiCoO2	0.0836664	T2	Electrolytes	0.05977
	Alpha-Fe2O3	0.0675474		High-voltage	0.0524314
	V2O5 films	0.0638271		Polymer	0.0504739
	Cathode	0.0617978		Thio-LISICON	0.0476543
	Process	0.0531336		Thin	0.0460922
	Li2S-Cu	0.0528221		Electrode	0.0454681
	Interfacial	0.0430099		Inorganic	0.0438584
	Particles	0.04245		Sulfur	0.0429712
	Composite	0.040384		Negative	0.0417726
	LixMn2O4 (LMO)	0.0388778			
T3	Films	0.01444608	T3	LLZO	0.04895287
	Model	0.01010294		Cubic	0.01721995
	Sputtering	0.00836155		Symmetry	0.01099688
	TiS2	0.00816412		Charge–discharge	0.00709187
	Magnetron	0.00786813		Li7La3Zr2O12	0.00702762
	Li/S	0.00712542		Sintered	0.00578623
	Discharge	0.00706965		Operated	0.00501226
	Li	0.00669281		Configuration	0.00481405
	Radio-frequency	0.00639552		Lithium-oxygen	0.00450383
	Conductivity	0.00639449		Pellet	0.00433608
T4	Electrolyte/electrode	0.0661156	T4	Amorphous	0.01889356
	Ionic	0.0520889		Thermal	0.01713587
	Weight	0.0425375		Organic	0.01496224
	Coulombic	0.0423535		Porous	0.01399585
	Electrolyte–electrode	0.0418752		Complex	0.01329573
	Contacts	0.0408151		Bis	0.01277152
	Efficiency	0.0384371		Mechanism	0.01250002
	Anions	0.0326497		SPE	0.01178497
	Doped	0.0266365		Crystal	0.01143879
	Polymer	0.02396816			
T5	Li7P3S11	0.02384236	T5	Hydride	0.02190056
	Carbon	0.02347841		CPE	0.01068218
	Milling	0.02173685		Crosslinked	0.00766991
	Change	0.02024504		Lithiation	0.0025338

Table 4 (continued)

Period	Top related technological terms	Probability	Period	Top related technological terms	Probability
	State	0.01973634		Galvanostatic	0.0022309
	Positive	0.01920095		Crosslinker	0.00177835
	Li7P2S8I	0.01463088		Peak	0.0012763
	Attention	0.01460873		Suitable	0.00114438

patent topics from the whole SAO structures. Tables 6 and 7 summarize several important semantic relationships based on scientific papers text mining and patent text mining with examples.

Technological subsystems generation

Firstly, the experts in ASSLIBs technology domain preview the topic clustering results. They divide the technology topics into three categories based on the research experience and battery components. The three categories also fully reflect the main development trends of ASSLIBs technology. The three categories of the technology are considered: *"Electrode"*, *"All-solid Electrolyte"* and *"Electrode–Electrolyte interface"*. The *"Electrode"* addresses the positive and negative electrode materials and advanced processes. *"All-solid Electrolyte"* addresses all types of solid electrolytes and their optimization. *"Electrode–Electrolyte interface"* addresses all technologies that can solve the poor conductivity in electrolyte layers. Then, the experts subdivide all topics and terms into the corresponding categories and layers based on the three categories that have been divided. Finally, the technical subsystems and corresponding topics of the ASSLIBs are generated, as shown in Table 8.

Table 8 shows the division of partial scientific topics and patent topics for each period. This division is useful for generating the technology evolution map, as it directly defines the vertical structure of the map of technology evolution.

Technology evolution map construction and analysis

It's understanding the path of the technology evolution fully that is significant for forecasting technology development trends. The map of the technology evolution based on the scientific papers and another based on patents are constructed in this step.

Firstly, according to the classification of the topic subsystem in Table 8, we construct the vertical axis of the evolution map. The vertical axis represents the technical subsystem obtained above, and the horizontal axis represents the different time periods. Then according to the relationships of topics in Tables 6 and 7, we finish the evolution map of ASSLIBs technology as shown in Figs. 4 and 5, and the different technical topics are distinguished in different colors in each period. Meanwhile, we use the yellow five-pointed star symbols to label the high frequency technical elements in the above SAO results in Figs. 4 and 5. From these related terms, we can understand the development process of ASSLIBs technology by analyzing the variation of these topics and their top related terms in each layer over time.

Table 5 Patents clustering results in ASSLIBs (partial)

Period	Top related technological terms	Probability	Period	Top related technological terms	Probability
T2	Thin-film	0.0179254	T2	Layer	0.01707653
	Electrode	0.01246927		Electrode	0.01122693
	State	0.01240681		Film	0.01097842
	Active	0.01089514		Sulfide	0.00959918
	Positive	0.00944897		Metal	0.00939479
	Depositing	0.00912123		Thin film	0.00929448
	Layer	0.00833819		Formed	0.00848201
	Solid	0.00828017		Conductive	0.00835518
				Mixture	0.00815523
				Positive	0.00813182
T3	Ceramic	0.04997856	T3	Layer	0.0268819
	Inside	0.02583325		Sulfide	0.02080412
	Disposed	0.01207603		Electrode	0.01820313
	Resistor	0.00944497		Solid	0.01635814
	Current	0.00862614		Laminated	0.01580886
	Sodium-sulfur	0.00739297		Positive	0.01567156
	Starting	0.00701056		Active	0.01540673
	Collector	0.0069396		Secondary	0.01438454
	Heating	0.00655391		Material	0.0139775
	Unit	0.00646518		Negative	0.01387571
T4	Resin	0.07360575	T4	Polymer	0.04828793
	Acid	0.0212703		Carbon	0.0427442
	Sensor	0.02118147		Region	0.02881005
	Lanthanum	0.01666047		Film	0.0264446
	Structure	0.01290744		Salt	0.0226161
	Lithium-air	0.01187816		Thin	0.02115384
	Elements	0.01105403		Connection	0.01917028
				Discharge	0.01543771
				Laminating	0.01485045
				Anode	0.01376082
T5	Phosphorus	0.04224068	T5	Crystalline	0.02173461
	Sulfide-based	0.03625816		Bonding	0.01724188
	Niobium	0.03530533		Ball	0.01183512
	Nitrogen	0.03501958		Crystal-grain-boundary	0.00989493
	Oxygen	0.03255137		Nonwoven	0.00260867
	Vanadium	0.02566078		Crystal	0.000868
	Tantalum	0.01885769		Surface-layer	0.00082674
	Sulfur	0.01704833		Angle	0.00078384
	Flexible	0.01694422		Transition	0.00065393
	Calcium	0.01348028			

Table 6 Important topics semantic relationships with examples based on scientific papers

No	Patent topics	SAO structures	The number of similar structure	Relationship
1	Polymer electrolytes	The addition of ceramics to the polymer electrolyte-enhanced the ionic conductivity	22/289	Enhance
2	Negative electrode active material	A new type of all solid-state lithium cell comprises a conductive nano-sulfur composite (NSC) cathode	15/289	Comprise
3	Solid electrolyte interface	The LiCoO(2) electrode-solid electrolyte interface decreased with the introduction of a NbO(2) interlayer at the interface	9/289	Decrease
4	Sulfide-based electrolytes	One is adding oxide to sulphide glasses (in oxysulphide systems)	18/389	Add
5	Positive electrode	Cathode materials include crystalline V2O5 thin films for rechargeable all-solid-state thin-film lithium microbatteries	11/289	Include
6	Conductive additive	MWCNTs are shown to be a superior conducting additive for Li batteries with a solid-state construction	7/289	Include
7	Garnet-type (LLZO)	Li/Al2O3-added LLZ/LiCoO2 configuration was fabricated	21/289	Add
8	NASICON-type	Ultrapure NASICON-type phase is synthesized by a 1,2-propylene glycol (1,2-PG)-assisted sol–gel method	20/289	Be synthesized by
9	Argyrodite-type	We prepare argyrodite-type Li6PS5X (X = Cl, Br, I) using mechanical milling followed by annealing	24/289	Prepare
10	Perovskite-type	Experimental results have also confirmed that nitrogen doping has improved ionic conductivity of LLTO	13/289	Improve
11	Binder	This large contact resistance is reduced by introducing an adhesive layer comprised of a mixture of LiBH4 and LiNH2	6/289	Introduce
12	Electrolytes modifier	Metal–organic framework (MIL-53(Al)) is a novel structural modifier for solid polymer electrolytes	9/289	Include
13	New electrode materials	The favorable electrochemical properties of TiH2 electrode which are superior to the previous report by means of conventional organic liquid electrolyte system	5/289	Be superior to
14	Metal hydride-based	Several studies have demonstrated that MgH2 is a promising conversion-type anode toward Li	7/289	Include
15	Sheet-style anodes	FeS nanosheets with thickness of 10 tun are synthesized by a poly (vinyl alcohol)-assisted precipitation method	13/289	Synthesize

Table 7 Important topics semantic relationships with examples based on patents

No	Scientific topics	SAO structures	The number of similar structure	Relationship
JP2009164059-A	Positive electrode	The contact area with constituent material of positive electrode increases with net-like conductor	15/378	Increase
JP2009187911-A	Solid electrolyte	Solid electrolyte formed in chain shape or single or multi-dimension mesh shape	20/378	Form
CN101557017-A	Electrolyte and electrode material	The main chain containing sulfur improves compatibility between electrolyte and electrode material	26/378	Improve
JP2010262764-A	Binder resin	The increase in electrical resistance of positive mix layer is suppressed by using styrene-containing binder resin	33/378	Is suppressed
US2010047696-A1	Ceramic material	The inclusion of aluminum enables the ceramic material to be sintered compact with increased degree of sintering density and conductivity	19/378	Improve
JP2012099323-A	Positive electrode active material	The surface of strong base positive electrode active material is covered by the coating layer provided with acidic polyanion structure	16/378	Is covered
WO2013024537-A1	The lithium-ion conductor	The lithium-ion conductor reduced internal resistance	19/378	Reduce
JP2014229495-A	Positive electrode active material	A positive electrode active material comprises zirconium-containing lithium nickelate (I) and lithium-zirconium oxide	12/378	Comprise
WO2014068777-A1	The electrical power collector	The electrical power collector has holes of three-dimensional meshwork structure	17/378	have
WO2015163152-A1	Garnet-type	Garnet-type contains zirconium carbonate complex	19/378	Contain
WO2015029289-A1	The solid electrolyte layer	The solid electrolyte layer has tetragonal crystal perovskite-type crystal structure	7/378	Have
JP2019186212-A	Negative electrode	Negative electrode contains silicon, and an electroconductive material	11/378	Contain
JP2019003901-A	Negative electrode	Negative electrode comprises porous single silicon material having specified density	18/378	Comprise
KR2020000192-A	Organic–inorganic composite	Organic–inorganic composite solid electrolyte comprises LixLayZrzMwO12	6/378	Comprise
CN108808109-A	Preparing method	Preparing method involves preparing a lithium niobium titanium oxysol solution by a sol–gel method	19/378	Involve

Table 8 Partial topics division in the technological subsystem

Subsystem	Scientific topics (time period)	Patent topics (time period)
All-solid electrolyte polymer	Polymer electrolyte (T1), thin-film structure (T1), sulfide-based electrolytes (T2), garnet-type (LLZO) (T3), composite polymer electrolyte (T5)	Thin-films structure(T2), polymer electrolytes(T2), sulfide-type electrolytes(T2), glass–ceramic electrolyte(T3), perovskite-type(T4)
Electrode	Negative electrode active material (T1), active electrode material (T2), oxide-based positive electrode (T3), nano porous carbon electrode (T5)	Oxide-based positive electrode (T3), negative electrode active material (T4), positive electrode(T4), electrode collector layer (T5)
Electrode–electrolyte interface	conductive additive (T3), electrolytes modifier (T4), binder(T4)	conductive additive(T3), lithium ion conductor (T4), binder(T5)

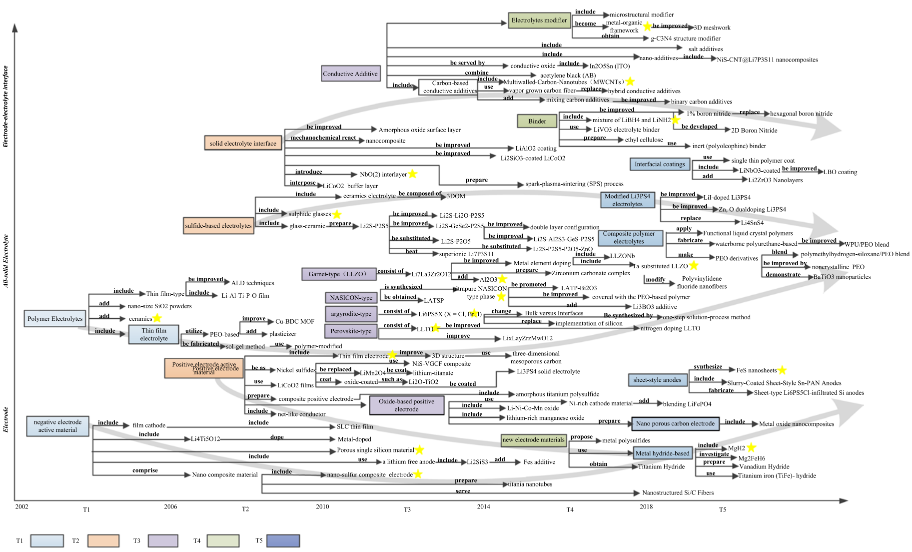

Fig. 4 Evolution map of ASSLIBs technology based on scientific papers

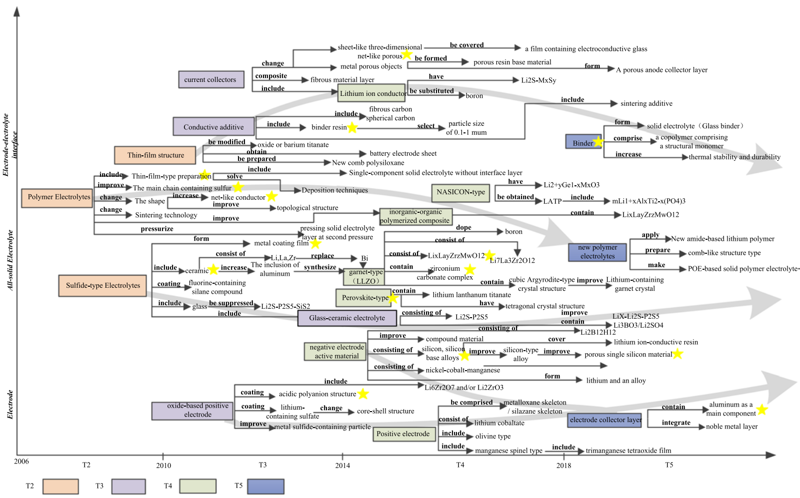

Fig. 5 Evolution map of ASSLIBs technology based on patents

In Fig. 4, in the electrode layer, there are mainly three technical topics including the negative electrode active material, positive electrode, and the new electrode materials. Top related technical terms such as metal-doped, Fes additive, Nanocomposite material, Nanostructured Si/C fibers have appeared with changes over time. That shows the negative electrode active materials are mainly divided into two evolution paths: adding metal and designing Nanostructure. Adding metal to form a protective layer can reduce the generation of dendrites and dead lithium in the cycle of the battery. As one of the most mainstream modification methods of the negative electrode material, it has been widely studied. Compared to the addition of metals, the design of nanostructures has been developed later, with ongoing research beginning around 2010. Designing and manufacturing the Nan micro structure such as the Nanopore array can change the surface of the electrode, to improve the overall performance. Top related technical terms such as LiMn2O4, Nis-VGCF composite, LiCoO2, Ni-rich cathode material have appeared with changes over time, which shows that the application of positive electrode materials have changed from LiFePO4 in 2010 to the Ni-rich, Li-rich, sulfur and so on. Compared to LiFePO4, Ni-rich, Li-rich and sulfur have higher energy density. Therefore, the combination of lithium metal with high load and high specific energy cathode material is considered to be a reliable way to improve the energy density of ASSLIBs. Top related technical terms such as metal polysulfides, Titanium hydride, MgH2, TiFe-hydride have appeared with changes over time, showing that metal hydride has gradually become a new electrode material for ASSLIBs. MgH2 with a tetragonal TiO2 structure has attracted much attention because of its good dynamic performance at high grow rate discharge since 2008. We predict that the clean, pollution-free, and safe materials may become future development opportunities, and the electrode materials will be multi-doped and more economical. The electrode microstructure design such as 3D structure and metal protective film will be paid more attention to in the future.

In the all-solid electrolyte layer, there are eight types of solid electrolytes related to ASSLIBs. The polymer electrolytes appeared in 2002, the sulfide-based electrolytes appeared later and the solid electrolyte based on oxide including the garnet-type, NASICON-type and Perovskite-type appeared around 2010. Currently, there are a variety of all-solid electrolytes, and the coexistence of multiple structures will continue to be a trend of the technology in the short term. In oxide-based electrolytes, LLZO is considered an ideal electrolyte material due to its high electrical conductivity and stability. Compared with oxide-based electrolytes, sulfide solid electrolyte also has been paid attention to since 2006 for its high ionic conductivity, low grain boundary resistance and high oxidation potential. In recent years, the organic–inorganic composite solid electrolyte has been more focused on, which combines the advantages of inorganic electrolyte and polymer electrolyte. This shows that the current research trend has shifted from polymer electrolyte and oxide electrolyte to organic–inorganic composite solid electrolyte. We predict that the development of organic–inorganic composite solid electrolytes will become a research focus in the future. At the same time, LLZO-type electrolyte interface manufacturing method will also receive continuous attention.

In the Electrode–electrolyte interface layer, the methods to improve the electrode–electrolyte interface mainly include adding a coating layer and adding conductive additives. Since 2010, researchers have modified the interface by introducing additives or modifiers to effectively reduce the interface resistance. In recent years, researchers have begun to focus on the inhibition of lithium dendrite growth by coating electrolytes. Top related terms such as single thin polymer coat, LiNbO3-coated and Li2ZrO3 nanolayers have appeared with changes over time, which show that researchers pay more attention to reducing the interface resistance by coating protective layer.

In Fig. 5, the research on the application of ASSLIBs technology has been widespread since 2006, a period later than the theoretical research of ASSLIBs. From the emergence of the new topics and top related terms in different layers and different periods, we can see that the applied research moves from electrode materials to the all-solid electrolytes and then to the electrode–electrolyte interface design. From this, we are able to understand the development process of ASSLIBs by analyzing the variation of technical terms in each layer over time.

In Fig. 3, in the electrode level, there are mainly about the positive electrode materials, negative electrode active materials and the electrode collector layer. Top related terms such as oxide-based positive electrode, core–shell structure and metal siloxane skeleton/silazane skeleton have appeared with changes over time, which shows that the research of positive electrode materials has gradually developed from oxide coating in 2010 to metal oxide particle doping in 2014, and then to the application of nanocomposite electrodes in 2018. It has great potential to use oxides as cathode coating targets to improve battery cycle performance. Moreover, the carbon-based nanocomposites can not only improve the electrochemical performance, but also improve the stability of the battery structure. From the shifting of the related terms, we can see that the design and development of high ionic conductivity, stable and safe cathode materials are the future research trends, and the nanocomposite electrodes will get more researchers' attention. Top related terms such as compound material and silicon-type alloy have appeared with changes over time, which shows that the negative material gradually changes from lithium metal to composite material and lithium alloy. For example, lithium silicon alloys appearing in 2016 are considered to be an alternative to lithium anodes. The shifting of electrode materials indicates that the materials with high stability and high conductivity are the future development trend. The organic–inorganic composite materials and alloy materials are a direction for future development.

In the all-solid electrolyte layer, there are seven types of structures related to the electrolytes. The polymer electrolyte and the sulfide-based electrode have developed since 2006. Since 2014, the oxide-based electrolyte such as LLZO-type and NASICON type has been widely researched. These results show that polymer electrolyte, sulfide electrolyte and oxide electrolyte are the three most important electrolytes in current research. New polymer electrolytes appeared in 2018, indicating that researchers began to focus on coating and modification of the traditional polymer electrolytes. A variety of all-solid-state electrolyte structures indicates that the coexistence of multiple structures will continue to be the trend of the technology in the short term.

In the Electrode–electrolyte interface layer, there are mainly the additive and the modifier to improve the conductivity of the electrode–electrolyte interface. The solid electrolyte film appeared in 2010, and the additive and binder appeared later. This development indicates that researchers began to focus on improving the overall performance of the battery by adding functional additives. Compared with adding film coating, the application of additives is more convenient and has comprehensive advantages. Therefore, the electrode–electrolyte interface development trend will involve a composite additive with comprehensive advantages in capacity, stability and conductivity.

Technology evolution trends analysis

Gaps analysis between scientific papers and patents

As shown in Figs. 4 and 5, there are evident differences between the evolution path based on scientific papers and patents. This indicates the research and development of the ASSLIBs technology in scientific papers and patents are also diverse. Based on the comprehensive research on the similarities and differences between scientific research and technology (Qi et al., 2018; Wang et al., 2015), we compare the technical topic clusters in T5 period, as shown in Table 10. In Table 10, the appearance time in papers and the

Table 9 Comparison of technology trends analysis based on scientific papers and patents

Topic clusters	Top related terms	The appearance time in papers	The appearance time in patents	Time lag
Polymer electrolytes	Thin film-type, PEO-based, Inorganic–organic Polymerized composite	2018	2019	1
Argyrodite-type	Li6PS5X(X = Cl,Br,I)	2018	2018	0
Garnet-type (LLZO)	LLZONb, Nanofibers, Cubic argyrodite-type	2018	2018	0
New polymer electrolytes	WPU, PEO derivatives	2018	2018	0
Solid electrolyte interface	Nanocomposite, Coatings, Film	2018	2019	1
Binder	Inert binder, Boron nitride, Copolymer	2019	2020	1
Positive electrode	Nano-metal oxide, Ni-rich	2019	2020	1
Conductive additive	Carbon-additive, Sintering-additive	2019	2019	0
Perovskite-type	Nitrogen doping LLTO Tetragonal crystal structure	2019	2019	0
Electrolytes modifier	Metal–organic, Microstructure	2020	2021	1
Interfacial coatings	LBO coating, Nano layers	2020	2021	1
Modified Li3PS4 electrolytes	Zn, O dual doping, LiI-doped	2021	–	–
Sheet-style anodes	Thin film	2021	–	–
Metal hydride-based	MgH2	2021	–	–
Electrode collector layer	Noble metal layer	–	2020	–
Current collectors	Lithium ion conductor porous Resin base	–	2020	–

appearance time in patents mainly refer to the earliest appearance of the technical topics in the T5 phase, which can be obtained from the patent publication data.

Due to the large period of the evolution paths, the relationships among technological topics have become increasingly complex. To analyze the latest development trend of ASSLIBs technology, this paper selects the topic clusters in the T5 period (2018–2021), which the topic clusters are the newest and the most comprehensive. As shown in Table 9, the Time lag referred to the time when the top related terms in clusters first appear. The visualization results are shown in Fig. 6.

In Fig. 6, every topic cluster includes related topics and the top related terms, represented by the dotted circles and the top related terms are expressed in solid circles. The correspondence relationships of the topic clusters are connected by dotted lines. In particular, red dotted lines represent potential technological opportunities.

As shown in Table 9 and Fig. 6, from 2018 to 2021, the topic clusters based on scientific papers were always ahead of the clusters based on patents, but the time lag was basically about one year and even less. Therefore, the topics which was only appeared in the scientific papers topic clusters or only appeared in the patents also could be considered as potential technological opportunities in the future. Specifically, the cluster Current collectors and the Electrode collector layer are the field in which technological developments appear but lack scientific research, also presenting a gap. This shows that these two topic clusters are widely used in practical technology R & D, but lack the supplement of scientific research. Therefore, these two topics can be regarded as technical opportunities in the field of scientific research after 2020. The topic cluster Electrolytes modifier appeared in scientific papers in 2020, and the Additive with 3D microstructure was first put forward to improve interface conductivity. However, there is no research on the additive of 3D microstructure that emerged in 2020 in patents, which also can be regarded as technological opportunities after 2020. In 2021, new topic clusters appeared in scientific papers, such as metal hydride-based, sheet-style anodes and modified Li3PS4, while there is none in patents. Therefore, these three topics can be regarded as technological opportunities after 2021.

Potential technical topics identification

In this stage, we calculate the growth rates and the intensity proportion of technical topics obtained from the scientific papers and patents and tried to find the topics with development potential.

Similarly, we select the technical topics in T5 period (2018–2021) including 13 topics based on scientific papers and 12 topics based on patents. The growth rate and intensity

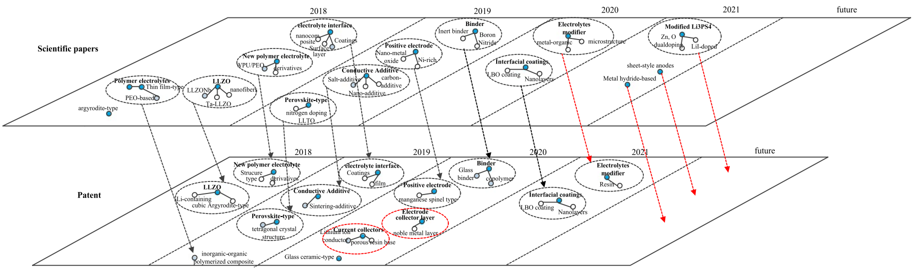

Fig.6 Technology trends analysis based on the gaps between scientific papers and patents

Table 10 The growth rate and the intensity proportion of the 25 topics

Quadrant	Topics based on scientific papers	Intensity proportion	Growth rate	Topics based on patents	Intensity proportion	Growth rate
I (High intensity proportion and High growth rate)	Conductive additive	0.141	1.56	Positive electrode	0.365	1.362
	Metal hydride-based	0.106	0.91	Glass ceramic-type	0.309	1.125
	Perovskite-type	0.23	0.7	Perovskite-type	0.175	0.662
	Interfacial coatings	0.283	1.308	Inorganic–organic polymerized composite	0.173	1.267
	New polymer electrolytes	0.354	0.853	New polymer electrolytes	0.117	0.779
II (Low intensity proportion and High growth rate)	Electrolytes modifier	0.303	1.162	Electrode collector layer	0.047	1.05
	Modified Li3PS4 electrolytes	0.048	0.842			
III (High intensity proportion and Low growth rate)	Positive electrode	0.147	0.448	Modified binder	0.127	0.5
	Sheet-style anodes	0.156	0.209	Electrolyte interface	0.234	0.411
	Solid electrolyte interface	0.25	0.328	Current collectors	0.3	0.5
IV (Low intensity proportion and Low growth rate)	Argyrodite-type	0.04	0.45	Conductive additive	0.019	0.333
	Binder	0.025	0.2	Garnet-type (LLZO)	0.062	0.182
	Garnet-type (LLZO)	0.089	0.146			

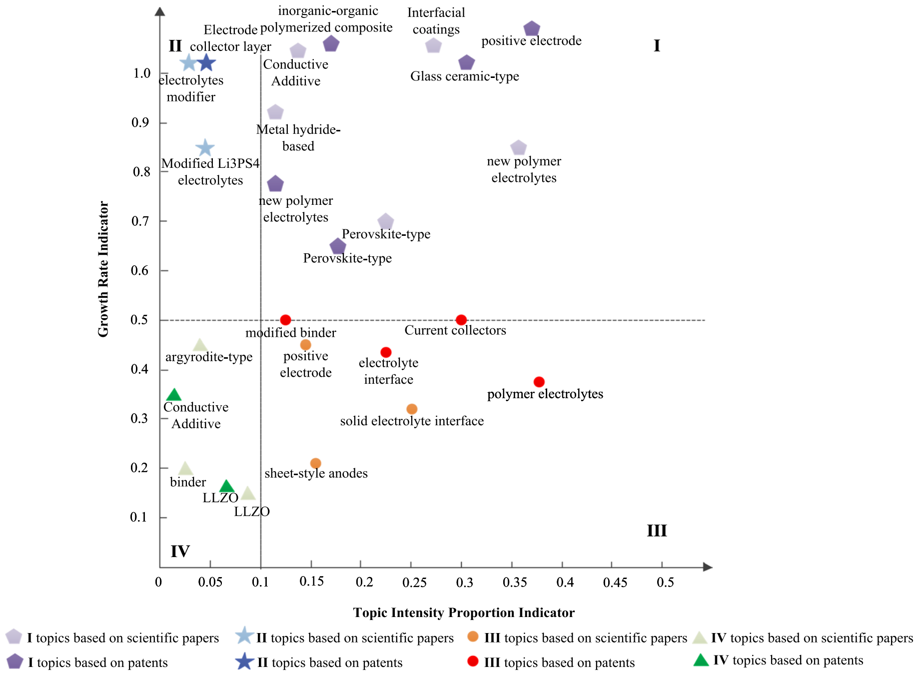

Fig. 7 The scatter diagram of the topics of ASSLIBs

proportion are computed for the 25 topics from two aspects, as shown in Table 10. Then the two-dimensional evaluation system is established with the growth rate and intensity proportion values, where the 25 topics were divided into four quadrants, as shown in Fig. 7.

As shown in Table 10 and Fig. 7, conductive additive, metal hydride-based, perovskite-type, new polymer electrolytes and the interfacial coatings in the scientific papers are the most important topics with the highest growth rate. Identically, the perovskite-type and the new polymer electrolytes in patents also keep the high intensity proportion and high growth rate. Based on these results, we could see that the corresponding topics in the scientific papers and the patents follow the same trend. Another seven topics are identified with relatively high intensity proportion shown in Quadrant II, reflecting the future development opportunities of ASSLIBs to a degree.

Based on the above analysis, we can predict the ASSLIBs technology development trends by combing the evolution maps and the scatter diagram based on the scientific papers and the patents. Therefore, as shown in Figs. 4, 5, 6, 7 and Tables 9, 10, the final technology trends and opportunities could be obtained, and the results are as follows.

(1) Perovskite-type electrolytes have emerged since 2010 and have been widely used in practical production. This topic shows a high growth rate and high theme intensity both in scientific papers and patents. Particularly, the non-metallic modification of LLTO will receive more attention in the future, which can be an innovation opportunity.
(2) The new polymer electrolyte came into public view after 2018. Similarly, this topic also shows a high growth rate and high intensity in scientific papers and patents, and

the growth rate and intensity in patents are higher. According to the above analysis, we predicted that the new polymer mainly composed of PEO derivatives will have higher growth potential after 2021.

(3) Interfacial coating technology first appeared in 2018, showing the highest topic growth rate and topic intensity in scientific papers. In recent years, the method of reducing the resistance between battery and electrolyte by interfacial coating has been widely recognized. Therefore, we can believe that this topic will have a higher growth potential in scientific papers after 2021.

(4) Glass–ceramic electrolytes, as one of garnet-type electrolytes, are widely used in the preparation of solid electrolytes, which also shows the highest growth rate and the highest intensity in patents. In recent years, the modification technologies for glass–ceramic electrolytes have been gradually enriched. Therefore, we can predict that it will have a higher growth rate and topic intensity after 2021.

Conclusion

The exploration of the technological evolution paths and identification of the technology innovation opportunities can be crucial to stimulate the technology innovation of enterprises. However, existing studies oriented to the technological evolution paths seldom consider the semantic linkages of technical terms. At the same time, current researches related to technology evolution lack a more in-depth and complete exploration of the technology development trends. In this paper, an integrated method including the LDA topic modeling, SAO semantic analysis and technology evolution mappings has been introduced. The multiple methods are combined to quantify the path information, which improves the accuracy of technical terms descriptions, and provides a new perspective for technology evolution research. The gaps analysis between science and technology is used to forecast technology development trends, and the growth rate and intensity proportion indicator are calculated to obtain potential technology topics. All-solid-state lithium batteries technology is selected as a case study, through which the proposed framework is proven to be valid and flexible. This study contributes to the technology forecasting methodology and sheds light on the emergence and future trends of technology studies.

However, there also exist some limitations that need to be improved in the future. Firstly, for the topic clustering method, we use only the LDA topic modeling to cluster technical topics. Although the LDA modeling can be used to cluster the documents which have a similar topic, it is not enough to capture the whole valuable topics. The existence of meaningless topics not only increases the computational workload, but also affects the accurate process of the technology evolution paths. In the future, the algorithm improvement and parameter tuning will be applied to further improve the accuracy of technical topic extraction. Secondly, comparing the gaps between topics based on scientific papers and patents is a meaningful method to predict technology trends. Although the gap analysis in this paper can be used to analyze the time lag of similar topics, it is not enough to justify the differences in aspects of the topics. The reason is that the differences in content and context of similar topics may also affect future trends. In the future, more quantitative analysis methods will be applied to calculate the deeper relationship between similar topics, which may be more comprehensive for forecasting technology trends. In addition, it is still not sufficient to consider only scientific papers and patents as the research data, which needs to be supplemented by multiple data.

Acknowledgements This research was supported by Henan Xing Culture Project Cultural Research Special Project (No. 2022XWH082); Henan Province universities key research project funding program(No.24A630033); Henan Province Soft Science Project supported by Department of Science and Technology of Henan Province; National Natural Science Foundation of China (No. 62173253).

Author contributions All authors contributed to the study conception and design. Related Work, refinement of the manuscript framework and post revision were completed by Lecturer Peng Liu. The first draft of the manuscript was written by Wei Zhou. Research methodology, and resources were performed by Prof. Lijie Feng. The supervision and funding acquisition were performed by Prof. Jinfeng Wang. The methodology and writing—review and Editing were written by Prof. Kuo-yi Lin. The algorithm design and optimization were completed by Xuan Wu. The Empirical data and results analysis were searched by Dingtang Zhang. All authors commented on previous versions of the manuscript and all authors read and approved the final manuscript.

Funding This research was supported by Henan Xing Culture Project Cultural Research Special Project (No. 2022XWH082); Henan Province universities key research project funding program(No.24A630033); Henan Province Soft Science Project supported by Department of Science and Technology of Henan Province; National Natural Science Foundation of China (No. 62173253).

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval I certify that this manuscript is original and has not been published and will not be submitted elsewhere for publication while being considered by *Scientometrics*. And the study is not split up into several parts to increase the number of submissions and submitted to various journals or to one journal over time. No data have been fabricated or manipulated (including images) to support your conclusions. No data, text, or theories by others are presented as if they were our own. The submission has been received explicitly from all co-authors. And authors whose names appear on the submission have contributed sufficiently to the scientific work and therefore share collective responsibility and accountability for the results.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Ba, Z., & Liang, Z. (2021). A novel approach to measuring science-technology linkage: From the perspective of knowledge network coupling. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2021.101167

Balland, P. A., & Boschma, R. (2022). Do scientific capabilities in specific domains matter for technological diversification in European regions? *Research Policy, 51*(10), 12. https://doi.org/10.1016/j.respol.2022.104594

Chen, M. C., & Ho, P. H. (2021). Exploring technology opportunities and evolution of IoT-related logistics services with text mining. *Complex & Intelligent Systems, 7*(5), 2577–2595. https://doi.org/10.1007/s40747-021-00453-3

Choi, S., Yoon, J., Kim, K., Lee, J. Y., & Kim, C. H. (2011). SAO network analysis of patents for technology trends identification: A case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells. *Scientometrics, 88*(3), 863–883. https://doi.org/10.1007/s11192-011-0420-z

Coccia, M. (2019). The theory of technological parasitism for the measurement of the evolution of technology and technological forecasting. *Technological Forecasting and Social Change, 141*, 289–304. https://doi.org/10.1016/j.techfore.2018.12.012

Feng, L. J., Liu, K. H., Wang, J. F., Lin, K. Y., Zhang, K., & Zhang, L. Y. (2023). Technology opportunity discovery of proton exchange membrane fuel cell from the perspective of papers-patents correlation analysis. *Frontiers in Energy Research, 10*, 13. https://doi.org/10.3389/fenrg.2022.1035657

Fernandes, M. D., Bistritzki, V., Domingues, R. Z., Matencio, T., Rapini, M., & Sinisterra, R. D. (2020). Solid oxide fuel cell technology paths: National innovation system contributions from Japan and the United States. *Renewable and Sustainable Energy Reviews*. https://doi.org/10.1016/j.rser.2020.109879

Hajikhani, A., & Suominen, A. (2022). Mapping the sustainable development goals (SDGs) in science, technology and innovation: Application of machine learning in SDG-oriented artefact detection. *Scientometrics*. https://doi.org/10.1007/s11192-022-04358-x

Han, X., Zhu, D., Wang, X., Li, J., & Qiao, Y. (2021). Technology opportunity analysis: Combining SAO networks and link prediction. *IEEE Transactions on Engineering Management, 68*(5), 1288–1298. https://doi.org/10.1109/tem.2019.2939175

He, X. J., Meng, X., Dong, Y. B., & Wu, Y. Y. (2019). Demand identification model of potential technology based on SAO structure semantic analysis: The case of new energy and energy saving fields. *Technology in Society*. https://doi.org/10.1016/j.techsoc.2019.02.002

Hernandez-Quintanar, L., & Rodriguez-Salvador, M. (2019). Discovering new 3D bioprinting applications: Analyzing the case of optical tissue phantoms. *International Journal of Bioprinting, 5*(1), 11. https://doi.org/10.18063/IJB.v5i1.178

Huang, Y., Zhu, F., Porter, A. L., Zhang, Y., Zhu, D., & Guo, Y. (2021). Exploring technology evolution pathways to facilitate technology management: From a technology life cycle perspective. *IEEE Transactions on Engineering Management, 68*(5), 1347–1359. https://doi.org/10.1109/tem.2020.2966171

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2018). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications, 78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Jiao, J. L., Zhang, A. F., Zha, J. R., & Li, J. J. (2023). Technological opportunity identification of cement kiln co-processing based on the gap between science and technology. *Journal of Material Cycles and Waste Management, 25*(1), 407–420. https://doi.org/10.1007/s10163-022-01547-x

Ke, X., Wang, Y., Ren, G., & Yuan, C. (2020). Towards rational mechanical design of inorganic solid electrolytes for all-solid-state lithium ion batteries. *Energy Storage Materials, 26*, 313–324. https://doi.org/10.1016/j.ensm.2019.08.029

Kim, K., Park, K., & Lee, S. (2018). Investigating technology opportunities: The use of SAOx analysis. *Scientometrics, 118*(1), 45–70. https://doi.org/10.1007/s11192-018-2962-9

Kim, S., & Yoon, B. (2021). Patent infringement analysis using a text mining technique based on SAO structure. *Computers in Industry*. https://doi.org/10.1016/j.compind.2020.103379

Kim, S., Yoon, S., Raghavan, N., Le, N. T., & Park, H. (2021). Developmental trajectories in blockchain technology using patent-based knowledge network analysis. *IEEE Access, 9*, 44704–44717. https://doi.org/10.1109/access.2021.3066573

Kumar, V., Lai, K. K., Chang, Y. H., Bhatt, P. C., & Su, F. P. (2020). A structural analysis approach to identify technology innovation and evolution path: A case of m-payment technology ecosystem. *Journal of Knowledge Management, 25*(2), 477–499. https://doi.org/10.1108/jkm-01-2020-0080

La Quatra, M., Cagliero, L., & Baralis, E. (2020). Exploiting pivot words to classify and summarize discourse facets of scientific papers. *Scientometrics, 125*, 3139–3157. https://doi.org/10.1007/s11192-020-03532-3

Lai, K. K., Bhatt, P. C., Kumar, V., Chen, H. C., Chang, Y. H., & Su, F. P. (2021). Identifying the impact of patent family on the patent trajectory: A case of thin film solar cells technological trajectories. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2021.101143

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

Lee, J., Ko, N., Yoon, J., & Son, C. (2021). An approach for discovering firm-specific technology opportunities: Application of link prediction to F-term networks. *Technological Forecasting and Social Change*. https://doi.org/10.1016/j.techfore.2021.120746

Leon-Silva, S., Fernandez-Luqueno, F., Zayago-Lau, E., & Lopez-Valdez, F. (2020). Silver nanoparticles, research and development in Mexico: A bibliometric analysis. *Scientometrics, 123*(1), 31–49. https://doi.org/10.1007/s11192-020-03367-y

Li, X., Wu, Y. D., Cheng, H. L., Xie, Q. Q., & Daim, T. (2023). Identifying technology opportunity using SAO semantic mining and outlier detection method: A case of triboelectric nanogenerator technology. *Technological Forecasting and Social Change, 189*, 31. https://doi.org/10.1016/j.techfore.2023.122353

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

Li, Z., Zhang, C., Jia, S., & Zhang, J. (2020). Galex: Exploring the evolution and intersection of disciplines. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 1182–1192. https://doi.org/10.1109/TVCG.2019.2934667

Liu, H., Chen, Z., Tang, J., Zhou, Y., & Liu, S. (2020). Mapping the technology evolution path: A novel model for dynamic topic detection and tracking. *Scientometrics, 125*(3), 2043–2090. https://doi.org/10.1007/s11192-020-03700-5

Liu, X., Wang, X., Lyu, L., & Wang, Y. (2022a). Identifying disruptive technologies by integrating multi-source data. *Scientometrics*. https://doi.org/10.1007/s11192-022-04283-z

Liu, Y., Wang, G., Zhou, Y., & Liu, Y. (2022b). Advanced technology evolution pathways of nanogenerators: A novel framework based on multi-source data and knowledge graph. *Nanomaterials (basel)*. https://doi.org/10.3390/nano12050838

Miao, Z., Du, J., Dong, F., Liu, Y., & Wang, X. (2020). Identifying technology evolution pathways using topic variation detection based on patent data: A case study of 3D printing. *Futures*. https://doi.org/10.1016/j.futures.2020.102530

Qi, Y., Zhu, N., Zhai, Y., & Ding, Y. (2018). The mutually beneficial relationship of patents and scientific literature: Topic evolution in nanoscience. *Scientometrics, 115*(2), 893–911. https://doi.org/10.1007/s11192-018-2693-y

Ren, H., & Zhao, Y. (2021). Technology opportunity discovery based on constructing, evaluating, and searching knowledge networks. *Technovation*. https://doi.org/10.1016/j.technovation.2020.102196

Shen, Y. C., Wang, M. Y., & Yang, Y. C. (2020). Discovering the potential opportunities of scientific advancement and technological innovation: A case study of smart health monitoring technology. *Technological Forecasting and Social Change*. https://doi.org/10.1016/j.techfore.2020.120225

Shi, X., Cai, L. F., Li, Z. P., & Cui, Y. Q. (2022). Exploring technological solutions for onboard hydrogen storage systems through a heterogeneous knowledge network: From current state to future research opportunities. *Frontiers in Energy Research, 10*, 17. https://doi.org/10.3389/fenrg.2022.899245

Sugimoto, C. R., Li, D., Russell, T. G., Finlay, S. C., & Ding, Y. (2011). The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. *Journal of the American Society for Information Science and Technology, 62*(1), 185–204. https://doi.org/10.1002/asi.21435

Takano, Y., & Kajikawa, Y. (2019). Extracting commercialization opportunities of the Internet of Things: Measuring text similarity between papers and patents. *Technological Forecasting and Social Change, 138*, 45–68. https://doi.org/10.1016/j.techfore.2018.08.008

van der Pol, J., & Rameshkoumar, J.-P. (2017). The co-evolution of knowledge and collaboration networks: The role of the technology life-cycle. *Scientometrics, 114*(1), 307–323. https://doi.org/10.1007/s11192-017-2579-4

Viet, N. T., & Kravets, A. G. (2022). The new method for analyzing technology trends of smart energy asset performance management. *Energies, 15*(18), 26. https://doi.org/10.3390/en15186613

Wang, J., & Yang, N. (2019). Dynamics of collaboration network community and exploratory innovation: The moderation of knowledge networks. *Scientometrics, 121*(2), 1067–1084. https://doi.org/10.1007/s11192-019-03235-4

Wang, M. Y., Fang, S. C., & Chang, Y. H. (2015). Exploring technological opportunities by mining the gaps between science and technology: Microalgal biofuels. *Technological Forecasting and Social Change, 92*, 182–195. https://doi.org/10.1016/j.techfore.2014.07.008

Wang, X., Ma, P., Huang, Y., Guo, J., Zhu, D., Porter, A. L., & Wang, Z. (2017). Combining SAO semantic analysis and morphology analysis to identify technology opportunities. *Scientometrics, 111*(1), 3–24. https://doi.org/10.1007/s11192-017-2260-y

Xu, H., Winnink, J., Yue, Z., Liu, Z., & Yuan, G. (2020). Topic-linked innovation paths in science and technology. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2020.101014

Xu, S., Li, L., An, X., Hao, L., & Yang, G. (2021). An approach for detecting the commonality and specialty between scientific publications and patents. *Scientometrics, 126*(9), 7445–7475. https://doi.org/10.1007/s11192-021-04085-9

Yang, C., Zhu, D., Wang, X., Zhang, Y., Zhang, G., & Lu, J. (2017). Requirement-oriented core technological components' identification based on SAO analysis. *Scientometrics, 112*(3), 1229–1248. https://doi.org/10.1007/s11192-017-2444-5