Journal of Scientometric Research, 2023; 12(1):204-210.
https://www.jscires.org

Research Article

A Methodology for Strategic Selection of Priority Research Topics in Terms of Bibliometric Analysis

Jung-Chol Jo[1], Jin-Bom Jong[2], Vasileios Drakopoulos[3,*], Song-Il Ri[4]

[1]Department of Clinical Medicine, Nampho Medical University, Ezzat Shahrour, NORTH KOREA.
[2]Department of Informatics, Institute of Science Strategy, Pyongyang, NORTH KOREA.
[3]Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, GREECE.
[4]Department of Mathematics, University of Science, Pyongyang, NORTH KOREA.

ABSTRACT

To ensure the strategic focus on science and technology development, it is essential to scrutinise the world research trends and demand as well as to strategically decide on priority research topics. This article suggests a methodological procedure to select strategically priority research and development subjects in terms of bibliometric analysis of the scientific literature database. Among the research topics studied frequently in a research field, we identify core subject topics with high multi-disciplinarity and give them priorities as strategic alternatives for research and development. The analytic procedure suggested can be applied effectively to discover the strategic options for research and development in any scientific area of research and development bodies that exercise the tracking type of research and development strategy.

Keywords: Bibliometric analysis; Science mapping; Subject network; Research popularity; Priority topic.

Correspondence:
Vasileios Drakopoulos
Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, GREECE.
Email: vdrakop@uth.gr
ORCID: 0000-0002-6478-3943

Received: 13-06-2022;
Revised: 21-12-2022;
Accepted: 15-02-2023.

INTRODUCTION

Given that science and technology have rapidly developed different fields of science linked with each other in complexity and diversity, the importance of strategic planning is getting more emphasized to decide the development direction of science and education in conformity with main world trends and national strategic goals as well as to raise the efficiency of the research and development, or R&D for short, through the choice and focus.

In R&D strategy formulation, the forecasting technique, that has been widely used so far, is the expert survey to integrate the different opinions of specific experts about the future prediction by processing statistically arbitrary estimations. It is advantageous on one hand while disadvantageous on other hand. Deciding the priority of R&D topics through expert survey or estimation has a high possibility of subjectivity intervention of some individuals. Moreover, requiring much cost and labour, the bibliometrics and science-map-based methods are recently prevailing and improving the quality of the strategic decision of expert-base.

Scientific publication, a core element in creating and spreading knowledge,[1] is a result of scientific research where scientists communicate their research results. The topic modelling technique, which analyses the occurrence of research topics representing specific research fields and their connection in scientific publication, makes it possible to detect the evolutional trends of research topics in scientific publications without requesting citation relationships.[2]

Mapping of Science based on co-word analysis on the scientific articles of "biotechnology and bioengineering," one of the key journals of the life sciences area, and analysed the research trend of scientific life area, was formulated in.[3] In[4], the topic network map based on a co-word analysis of more than 500,000 doctoral dissertations published from 1996 to 2014 was analysed. It made sure that neuroscience is becoming a key research topic as a bridge between different branches of health science. The references of literature in leading journals of library science or information science were analysed in,[5] where it is also verified the multi-disciplinarity of this research field in terms of subject classification, topic distribution and differences by discipline. In, [6] improved co-word analysis based on a conceptual matrix and semantic distance was suggested and showed that this analysis represents a much better result in matrix order and clustering results. In,[7] the newly occurred research fields and cross-disciplinary cooperation research based on co-word analysis on papers published in 1998-2007 of biological fuel fields of the WOS database were presented. It is concluded that many research institutions studying biofuel as a cross-discipline are

DOI: 10.5530/jscires.12.1.018

more concentrated on solving the problems arising in renewable energy field. It is stressed in[8] that connections through citation between patents in the nanotechnology field could be considered as a multilateral interaction index between science and technology. In[9] nanoscience and Nanotechnology were presented as activities classified as being associated with intermediation by measuring the cited connections between a research paper and a patent. In[10] the multi-disciplinary index expressed by the product of log values of term occurrence frequency is used whereas several cluster distributions and other measuring indices are based on term occurrence frequency.

In this article, a methodological procedure to discover the strategic R&D alternatives is proposed by classifications based upon identifying the core topics not only recently much studied but of high multi-disciplinarity in the field through bibliometric analysis on the database of scientific literature. Moreover, the dynamic evolution pattern of their research popularity was analysed and confirmed its usefulness by applying it to discover strategic alternatives in the field of nanoscience.

Identification of core topics

Degree of popularity

According to,[10] popular topics are terminologies that show high occurrence frequency in scientific literature. Degree of popularity, an index that represents the relative research strength of research topics, is calculated by summing the occurrence frequency $\varphi_{d,j}$ of topic j in literature d. Standardised degree of popularity NP is calculated as a ratio of the popularity of topic j by the sum of popularities of all topics in a discipline S

$$NP(j) = \sum_{d} \varphi_{d,j} \Big/ \sum_{j \in S} \sum_{d} \varphi_{d,j}.$$

In bibliometrics, standard degree of popularity represents occurrence probability of the research topic in a literature set. Representing the publishing year of a research literature d by $py(d)$, the standard annual degree of popularity of subject j is calculated as:

$$NP(j) = \sum_{d|py(d)=t} \varphi_{d,j} \Big/ \sum_{j} \sum_{d|py(d)=t} \varphi_{d,j}.$$

The standard annual degree of popularity represents the occurrence probability of a given research topic published in a year. Since there is a big difference between the number of articles in every field and the literature set, the standard research degree of popularity is used to compare relative research intensity correctly.

Core topics

In[11] the index expressed by the product of the log value of the occurrence frequency and the number of cluster distribution of the topic is used as a multi-disciplinarity criterion since the greater its value, the more massive and pivotal role it plays in the topic network. In fact, in a topic network, the node of high occurrence frequency, the hub node, stands for the research topics investigated much popularly, and the node of many cluster distributions plays an important role as a bridge that connects many disciplines in research topic networks.

In this article, we define research topics, not only connected with many disciplines but also of high occurrence frequency, as *core topics*, and then identify them by using the *core topic* index

$$CI(j) = Cd_j \times log\left(\sum_{d \in D} \varphi_{d,j}\right),$$

where $\varphi_{d,j}$ j is the occurrence frequency of topic j and Cd_j is number of cluster linked (co-occurred) by topic j in discipline D. Standardised core topic index is

$$NCI(j) = \frac{Cd_j \times log\left(\sum_{d \in D} \varphi_{d,j}\right)}{\sum_{j \in D}\left(Cd_j \times log\left(\sum_{d \in D} \varphi_{d,j}\right)\right)}.$$

The topics which have relatively high $NCI(j)$ value are the core research topics.

Identification of strategic alternatives based on evolutionary pattern analysis

Evolutionary pattern analysis by regression modelling

Since core research topics identified in the previous section represent topics in scientific literature published in the whole measuring period, the latest trend and evolution character is correctly unknown. Analysing dynamic changing characteristics with a time of degree of popularity of the core research topics with regression modelling, we can classify the evolution characters of these core topics into some groups.

For regression modelling, the time t is taken as an argument while the standardised annual degree of popularity of core research topics as the dependent variable. Polynomial regression model expressed by

$$Y = \alpha + \sum_{i=1}^{m} \beta_i\left(t^i\right) + \varepsilon$$

enables us to draw out the somewhat fitting curves of the degree of popularity for core topics.

To find the correct value of m, firstly, set the order of polynomial $m = 1$ and valid coefficient of correlation $R = 0.7$, using the least squares method. If all topics pass the test, choose the current order; if none of them passed, modify the model by increasing m. In this way, regression curves of degree of popularity for all core subject topics are drawn and classified into four categories according to their morphology; type A: "ascending", type B: "descending-ascending", type C: "ascending-descending", type D: "descending".

Life cycle analysis of evolution pattern and decision of strategic R&D alternative

These four evolution patterns can be explained as the "S curve" for the life cycle of science development. At the initial step of scientific research, new ideas were suggested continuously, new theories were added surrounding them and, in the meantime, related research topics that were developed, while papers increased quickly. Research topics follow their development tracks of the "S curve" until the topic research was completed and converted into some patent or radical questions were suggested.

At which step of the life cycle the research topic belongs can be estimated by the evolution pattern of its degree of popularity; if the annual degree of popularity of the research topic is continuously increasing, the topic can be considered at the development or growing step, whereas if it is decreasing, it will be at maturing or decline step. According to four evolution patterns, the core research topics would be classified into strategic R&D alternative.

High-technology strategy option

Research topics of pattern A are comparatively high-tech research subjects that appeared relatively late and continued to grow in research scale at the worldwide range. These are considered subjects that have global competence. This group's research topics can become strategic and innovative high-technology projects that would pull the future competitive industries.

Tracking strategy option

Research topics of pattern B enjoyed an upswing long ago but had exhibited descending trend in some periods for some reason and then have ascended again. These are research subjects of continuous development domain which have long development history and relatively enough knowledge stockpile. Although rather low in high technology content compared to topics of pattern A, they are considered R&D alternatives for tracking strategy options because development tracks can be discovered clearly in the science map.

Switchover strategy option

The topics of pattern C are regarded as research subjects that are matured enough and already converted into patents or came to decline owing to special reasons like ethical problems. These subjects have a lot of relevant literature and rich knowledge accumulation.

With these research topics, it is challenging to expect new scientific discoveries or radical innovations. They can be still selected as switchover research options making new innovative breakthroughs by changing the research orientation to new subjects relying on plenty of accumulated scientific knowledge.

Application Oriented Option

The research topics of pattern D are the subjects that have already come up to decline in the research life cycle with a long research history. So, they are research subjects where radical innovations are no longer expected. Such issues have already converted entirely or partly into patent technologies or for R&D strategy so that they can be selected as application-oriented projects like performance gap improvement or knowledge transfer.

Exploring the strategic alternatives in nanomaterial research field

Research structure in the field of nanoscience

For bibliometric analysis, an international scientific article database built in Sci-Tec Complex, DPRK was selected where as a science mapping tool, the software VOSviewer 1.6.10 was used. First, titles of all scientific articles published from 2000 to 2015 were searched for and eventually total of 279849 scientific articles connected with nanoscience and technology were retrieved.

Figure l shows the topic co-occurrence network map generated by using VOSviewer 1.6.10 for 279849 selected scientific papers in the nanoscience field.

In the map, each node stands for a topic and its size represents the occurrence frequency of the case, i.e., the number of articles related to given topics. The link between two nodes means that two issues co-occur in the same form and the thickness of the link, that is, the weight of the association, represents several articles that two topics co-occur. For each node in a map, node ID, the discriminant number of topics, is signed. It can be seen that 279,849 scientific articles are clustered into some coloured groups: red, blue, green, yellow, purple etc. The colour on each node represents the science branches; the red group at the left bottom of the figure stands for Electronics, a green group at the right top for Biology, the Blue in the right middle for Materials, yellow for Chemistry, purple for Engineering and pink does Basic research. Besides, there are small clusters.

Table l shows the research structure in terms of the relative size of the total number of topics, occurrence and co-occurrence by research branches of Nanoscience. As shown in Table l, in the field of nanoscience, nano-electronics is the largest research branch during 2000-2015; its total number of occurrence frequency of relevant topics is about 27.7 percent, next comes nano-biology with 16.1 percent, and then nano-material 15.12.1 percent.

As shown in Table 2, research subjects such as blend, clay nanocomposite, caprolactone, composite, nanofiber, carbon nanotube nanocomposite, etc., are conspicuously studied subjects in the nanomaterial cluster during 2000-2015.

Surveying corresponding papers in detail, many papers handle with diversity of organic compounds such as caprolactone,

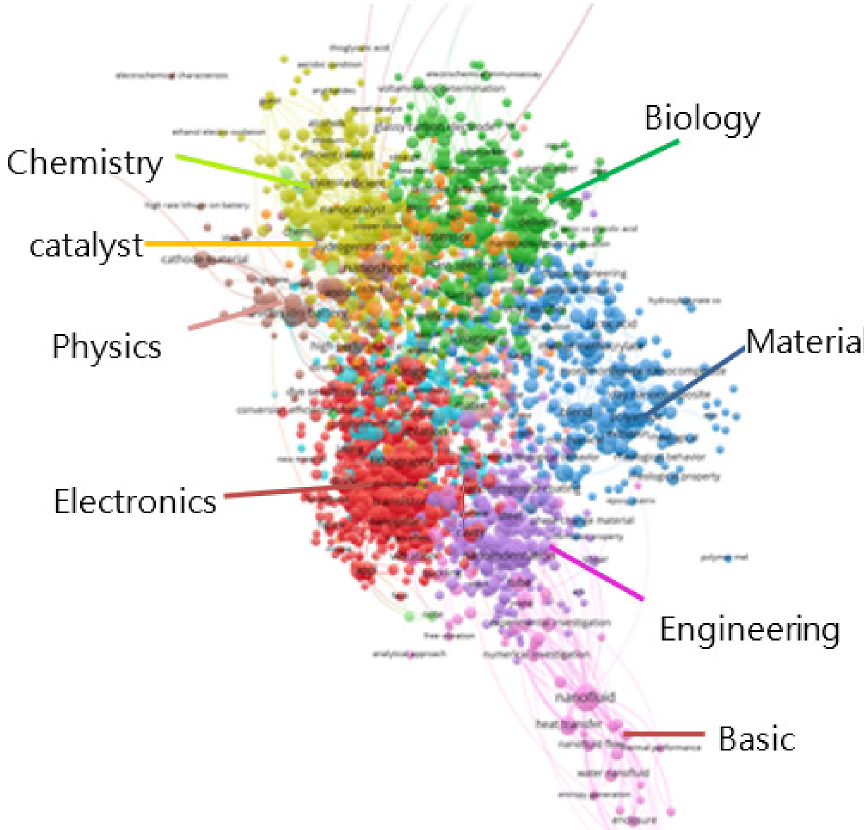

Figure 1: Topic co-occurrence network map in nanoscience field (2000-2015).

Table 1: Research structure in the field of nanoscience.

Cluster (Research branch)	Number of topics	%	Number of occurrence	%	Number of co-occurrence	%
1. Electronics	786	29.3	52888	27.7	33178	20.5
2. Biology	399	14.9	30712	16.1	30234	18.7
3. Material	302	11.3	23107	12.1	20338	12.6
4. Chemistry	301	11.2	21464	11.2	22108	13.7
5. Engineering	267	10.0	20371	10.7	16335	10.1
6. Basic	186	6.9	9431	4.9	5782	3.6
7. Catalyst	156	5.8	10017	5.2	7832	4.8
8. Physics	78	2.9	8334	4.4	7063	4.4
Other	205	7.6	14898			

polyimide, polyamide, polypropylene, epoxide, nylon, natural and compound rubber, etc., and excellent properties such as high heat stability, incombustibility, mechanical intensity of nanomineral compounds reinforced with different organic compounds fibers and nonorganic fibers such as glass or carbon.

These topics are research subjects with top popularity in nanomaterial branch, so they represent typical hub nodes in nanomaterial cluster.

The topics in Table 3 are not only research subjects which are mostly studied within the field of nanomaterial but also subjects which are handled a lot in different areas of Nanoscience such as nanoElectron, nanoBiology, nanoChemistry and nanoEngineering, etc., so they can be said to be basic and core nanoMaterial subjects that influence the whole of Nanoscience momentously.

Table 2: Research topics with top popularity in nanomaterial cluster.

ID	Subject topic	Number of occurrence	Number of co-occurrence	NP
194	Blend	573	601	0.00314
352	Clay nanocomposite	502	566	0.00296
232	Caprolactone	244	221	0.00116
393	Composite nanofiber	123	162	0.00085
261	Carbon nanotube nanocomposite	102	126	0.00066
119	Atom transfer radical polymerization	71	113	0.00059
432	Copolymerization	97	93	0.00049
166	Biocomposite	69	91	0.00048
201	Bone tissue engineering	104	91	0.00048
223	Calcium carbonate	75	90	0.00047
391	Compatibilizer	117	89	0.00047
272	Carbon nanotubes nanocomposite	41	44	0.00023

Table 3: Research topics with top core subject index in nanomaterial science.

ID	Subject topic	NP	NCI
194	Blend	0.00314	0.157491
352	Clay nanocomposite	0.00296	0.115625
400	Computational modeling	0.00012	0.072211
237	Carbon fiber composite	0.00013	0.058776
1074	Hydrogen sorption	0.00015	0.004331
432	Copolymerization	0.00049	0.031147
186	Bionanocomposite	0.00045	0.020592
2398	Superconductor	0.00025	0.011665
1602	Nanomagnet	0.00024	0.006744
183	Biomimetic synpaper	0.00021	0.01153
1446	Molecular simulation	0.000191	0.006212

Evolution patterns of core research subjects

For core subjects to be identified, dynamic changing characters on research popularities were examined using regression analysis. Figure 2 shows the standardised research popularity curves prepared by regression model for some core research subjects, ID183: biomimetic synthesis; ID1074: Hydrogen sorption; ID400: Computational modeling; ID186: Bionanocomposite.

As seen in the figure, changing characters of research popularity are different according to years and are roughly divided into four patterns. Table 4 shows dynamic evolution patterns of topic popularity of main core subjects in nanomaterial area. The patterns of most subjects are ascending (type *A*), and then descending-ascending (type *B*) comes next, while subjects of ascending-descending (type *C*) and descending (type *D*) are fewer.

Strategic R&D alternatives

According to evolution patterns identified, core research topics can be chosen as priority alternatives for R&D strategy.

High-technology strategy option

Topics that have evolution pattern A, such as biomimetic sandpaper, carbon nanotube composite, carbon fiber composite, clay nanocomposite, nanofiltration membrane, hydrogen sorption, superconductor, etc., can be selected as innovative research objects which can lead the future high-technology material industry.

Tracking strategy option

Research topics such as nanoceramics, drug carriers, nanocomposite coating, computational modelling, etc., can be selected for imitative or tracking strategy. They could get up to a

Table 4: Ynamic evolution patterns of top core subjects in Nanomaterial area.

Type	Tendency	Evolution pattern	Topics
A	Ascent	↗	ID183: Biomimetic synpaper ID263: Carbon nanotube composite ID237: Carbon fiber composite ID352: Clay nanocomposite ID394: Nanofiltration membrane ID432: Copolymerization ID:951: Graphene composite ID1446: Molecular simulation ID1073: Hydrogen sorption ID2398: Superconductor
B	Descent-ascent	↘ ↗	ID1509: Nanoceramic ID400: Computational modeling ID573: Drug carrier ID1514: Nanocomposite coating ID194: Blend
C	Ascent-descent	↗ ↘	ID186: Bionanocomposite ID324: Nanowhisker ID1602: Nanomagnet
D	Descent	↘	ID1935: Piezoelectric property ID1481: Nanocrystallization

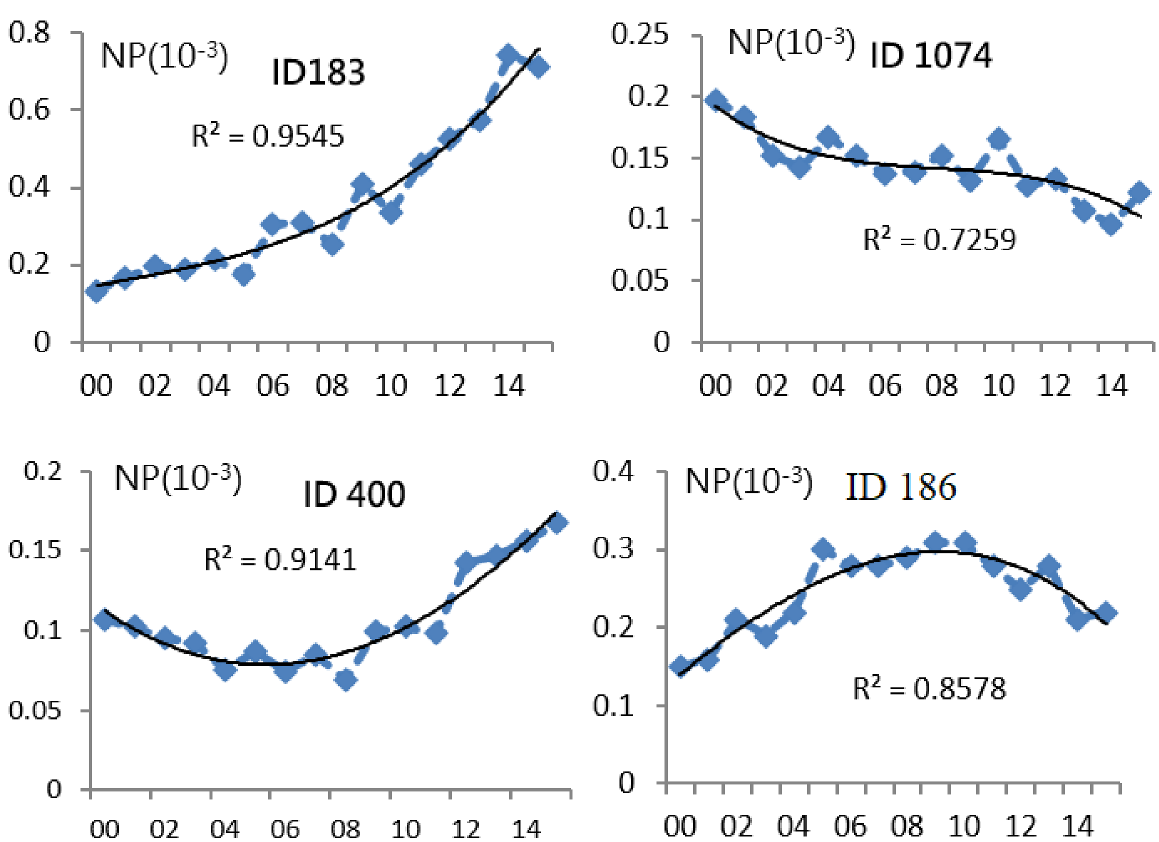

Figure 2: Research popularity curves for some core subject.

higher level quickly with actual feasibility since they have a long history and enough knowledge accumulation.

Switchover strategy option

Research topics such as bionanocomposite, nanowhisker, nanomagnet, etc., have already declined since research has matured enough or some life safety questions arose for nanotechnologies. Nevertheless, these subjects can be chosen as switchover strategic alternatives because they could be linked to innovative research breakthroughs by changing research direction to seek new topics referring to plenty of accumulated scientific knowledge.

Application oriented strategy option

Research topics such as piezoelectric, nanocrystallisation, etc., are the research themes that are matured enough and mostly converted into patent technology. These objects can be chosen as application-oriented research subjects for technological introduction or technology transfer issues.

CONCLUSION

In this article, a methodological procedure to identify research alternatives for R&D strategy on bibliometric analysis of a big database of science literature is suggested, based on analysing the research structure and the dynamic evolution characteristics of thematic topics in a particular field of science.

Through bibliometric analysis, the degree of research popularity and core topic index is defined in terms of occurrence and co-occurrence frequencies of different thematic topics. Then core topics that play an important role as the hub and pivot in the science field are identified and their dynamic evolution patterns are analyzed using a polynomial regression model. Eventually, according to the four types of evolution patterns, main R&D alternatives are classified by strategic options.

The proposed methodology was implemented to explore strategic research alternative topics in the field of nanomaterial science. The science mapping software VOSviewer 1.6.10 on the international science literature database of "Sci-Tech complex" was used and the research structure of the nanoscience field was analysed. Core topics of nanomaterial area were also identified. Finally, dynamic evolution patterns of core topics were analysed using the regression technique, and alternatives for nanomaterial research were classified into four strategic groups.

This methodology provides a practical tool that enables decision-makers to consolidate the subjective expert survey method bibliometrically by relying on the statistical integration of experts' arbitrary estimations in selecting strategic R&D alternatives. The procedure can also be applied effectively to exploring the strategic R&D topics in any scientific field of R&D bodies that exercise the tracking type of strategy.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Verbeek, A., Debackere, K., Luwel, M., Andries, P., Zimmermann, E., and Deleus, F. Linking science to technology: Using bibliographic references in patents to build linkage schemes. Scientometrics. 2002;54(3), 399-420.
2. Mei, Q., and Zhai, C. Discovering evolutionary subject patterns from text: An exploration of temporal text mining. In Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining. 2005;198-207. ACM
3. Nees Jan Van Eck, Methodological Advances in Bibliometric Mapping of Science, 2011:143-60.
4. Ran Xu, Navid Ghaffarzadegan. Neuroscience bridging scientific disciplines in health: Who builds the bridge, who pays for it?, Scientometrics. 2018;117:1183-204.
5. Zhang, G., Feng, Y., Yu, G., Liu, L., and Hao, Y. Analyzing the time delay between scientific research and technology patents based on the citation distribution model. Scientometrics. 2017;111:1-20.
6. Jia Feng, Yun Qiu Zhang, Hao Zhang. Improving the co-word analysis method based on semantic distance. Scientometrics. 2017;111:1521-31.
7. Janaina Gomes, Homero Dewes. Disciplinary dimensions and social relevance in the scientific communications on biofuels. Scientometrics. 2017;110:1173-89.
8. Meyer M. Does science push technology? Patents citing scientific literature. Research Policy. 2000;29(3): 409-34.
9. Meyer M. What is special about patent citations? Differences between scientific and patent citations. Scientometrics. 2000;49(1):93-123.
10. Griffiths TL, Steyvers M. Finding scientific topics. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(Suppl 1(1):5228- 35.
11. Haiyun Xu, Ting Guo, Zenghui Yue, Lijie Ru, Shu Fang. Interdisciplinary topics of information science: a study based on the terms interdisciplinarity index series. Scientometrics. 2016;106:583-601.

Cite this article: Jo JC, Jong JB, Drakopoulos V, Ri S-I. A Methodology for Selection of Strategic Priority Research Topics in Terms of Bibliometric Analysis. J Scientometric Res. 2023;12(1):204-10.