

Evaluation of Science and Technology Projects of Georgian Information System

Levon Chobanyan^{*}, Nelly Makhviladze^{**}, Pikria Tsotskolauri^{*},
Marina Lebedeva^{*}

^{*}Department of Information Technology Development, Institute Techniformi, Georgian Technical University

^{**}Department of Analysis and Information Resources, Institute Techniformi, Georgian Technical University

(Presented by Academy Member Ramaz Khurodze)

To solve the problems of evaluating and analysis of scientific and technical projects and making decisions at all stages of project development, it is necessary to create Georgian information systems that establish thematic relationships between science and technology. Such a system is being formed on the basis of a database of SciTech projects and international patent databases of Georgian Scientific and Technical Information Institute – Techniformi. For the analysis of productivity of scientific and technical projects, SciVal, a tool for citation-based research evaluation using Elsevier’s Scopus database, was used. It lets visualize research performance, benchmark relative to peers, develop collaborative partnerships and analyze research trends by using the cited patent data from scientific publications and other patent documents contained in the SCOPUS database. Based on the extended databases of Clarivate Analytics, Georgian Scientific and Technical Information Institute – Techniformi has developed a methodology for analyzing the relevance of a research project: determining the relevant IPC (International Patent Classification) indices, as well as the most cited patent documents, which make it possible to identify expediency of research. High-rated publications and the International Patent Classification indices are placed in the Techniformi DB of scientific and technical projects as fields of the bibliographic description of the corresponding scientific project. They are available to both project executors and science management teams. Depending on the results of the analysis, the project developers are provided with information about the project expediency, which is of a recommendatory nature. The tasks of the system for analyzing the relevance of SciTech projects developed at the Techniformi are similar to the tasks of the Derwent Data Analyzer (DDA) – a desktop data-mining platform that converts patent data, scientific literature and business intelligence into actionable commercial insights. © 2024 Bull. Georg. Natl. Acad. Sci.

information system, patent data, scientific publications, citation, relevant indexes

To solve the problems of monitoring and evaluating scientific and technical projects and making decisions at all stages of project development, it is

necessary to create Georgian information systems that establish thematic relationships between science and technology.

Science and Technology

In Georgian Scientific and Technical Information Institute Techinformi, such a system has been created on the basis of a database of scientific projects and international databases of patent information.

The importance of using patent information to establish thematic relationships between science and technology was explored [1].

The use of the established relationships for making decisions on the financing of scientific and technical projects was considered [2]. Based on a large-scale study of a decade-long array of the US and European patent descriptions, it was concluded that technological development is strongly dependent on the degree of funding scientific and technical projects.

Performance analysis can be done using the Scopus databases or the Clarivate Analytics: Web of Science, Derwent World Patents Index (DWPI) and Derwent Patents Citations Index (DPCI) databases.

For performance analysis using the SCOPUS system, SciVal is used, an online tool that offers access to research results using bibliometrics. It allows you to visualize research performance and analyze research trends.

Scopus databases contain over 43.7 million patent documents published by 5 patent offices:

1. World Intellectual Property Organization (WIPO)
2. European Patent Office (EPO)
3. United States Patent Office (USPTO)
4. Japan Patent Office (JPO)
5. UK Intellectual Property Office (IPO.GOV.UK)

Patents, Data Analysis

SciVal uses the citation data in patent documents from scientific publications as well as other patent documents contained in the SCOPUS database. This allows you to establish certain links between scientific developments and patent research, which indicates the practical results of the use of scientific research.

At the same time, it is extremely difficult to achieve high search rates for patent documents corresponding to these scientific and technical studies based on the SCOPUS database.

The main reason for this is the different principles for constructing the classification systems of patent and scientific information [3].

The International Patent Classification (IPC) is based on the functional-industry principle, while scientific classifiers are based on the subject area of research. These differences make it impossible to establish correspondence between the headings of scientific classifiers and the headings of the International Patent Classification.

To a large extent, these difficulties are overcome in the Clarivate Analytics: Web of Science, Derwent World Patents Index (DWPI) and Derwent Patents Citations Index (DPCI) databases. In the Derwent World Patents Index (DWPI) and Derwent Patents Citations Index (DPCI) patent databases, abstracts and keywords were created and entered, as a result of which it became possible to establish correspondence with abstracts and keywords from bibliographic descriptions of research papers.

The Project Evaluation Structure

Based on the extended databases of Clarivate Analytics, Georgian of Scientific and Technical Information Institute – Techinformi developed a methodology for analyzing the relevance of a research project: determining the relevant IPC (International Patent Classification) indices, as well as determining the most cited patent documents, which make it possible to determine promising continuations of research [4].

To determine the most pertinent IPC indices, the project explored the possibility of taking into account the average number of citations per document for different IPC headings and introducing an additional parameter of citation efficiency for IPC headings relevant to the search criteria [5].

High-ranking documents and indices of the International Classification of Inventions, obtained

as a result of the analysis are placed in the database of scientific and technical projects created in Georgian of Scientific and Technical Information Institute – Techinformi, as fields of the bibliographic description of the corresponding scientific project, and are available to both project executors and science management bodies.

At the first stage of the analysis, keywords are selected from the project abstract contained in the database of scientific and technical projects created in Georgian of Scientific and Technical Information Institute – Techinformi.

Based on them, an initial query is created to the Derwent World Patents Index (DWPI) database. Based on the information received, keywords are corrected: non-informative keywords with zero or an abnormally large number of relevant documents are removed, logical links between terms are established. As a result, the main query is formed, which corresponds to the list of IPC headings pertinent to the researched scientific and technical project [6].

At the second stage, the selection of patent documents belonging to each separate IPC heading, established at the first stage, is carried out, and time series of distribution by years of relevant

documents are determined for each IPC heading for the last 15 years. Citation data for each selected patent document is also stored.

Analyzing the time series, we can conclude that there is an increase or decrease trend in patent documents corresponding to the subject area of the researched scientific and technical project, and belonging to this IPC heading.

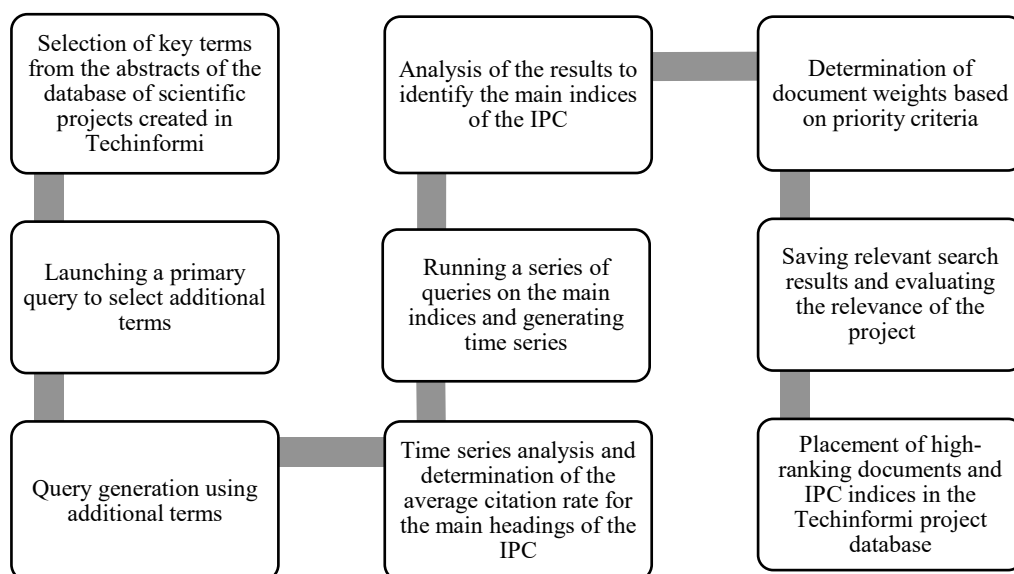
Depending on the results of the analysis, the project developers can be given information about the most priority technical areas for this scientific and technical subject. Of course, this information can only be of an advisory nature.

The results of the analysis can also be used by the Ministry of Education, Science and Youth of Georgia and/or the Shota Rustaveli National Science Foundation (SRNSF) when evaluating scientific and technical projects.

Below is the structure of the scientific project analysis procedure (Table).

At the first stage, the choice of keywords was made from the text of the project abstracts selected from the database of scientific and technical projects of Techinformi. Searching for selected keywords in the DWPI database yields dozens of

Table. The structure of the scientific project analysis procedure



IPC indices, thousands of patent documents, and the ability to select additional keywords.

Next, queries are generated with the addition of keywords selected from the DWPI database. At the same time, the number of selected IPC indices averages 3-6 rubrics. At the next stages, the distribution of documents of each rubric by years is determined, and time series are generated for each rubric. In addition, the average citation rates of patent documents for each of the main headings are determined.

After that, the relevance of the project is assessed and high-ranking documents and IPC indices are placed in the Techinformi project database.

The developed analysis procedure was tested on 10 research projects of the Institute of Cybernetics and 4 projects of the Institute of Metallurgy and Materials Science. The implementation of the first 4 blocks of the procedure revealed several (from 3 to 7 IPC indices), and several hundred patent documents corresponding to these indices. Further, when performing blocks 5, 6, 7 of the analysis procedure, time series of publication activity relevant to the draft patent documents were obtained for each IPC heading, covering the period from 2004 to 2021. When determining the pertinence of IPC headings and patent documents,

preference was given to headings with increasing activity over time, and patent documents with a large citation index and a later publication date. In addition, as a result of the analysis, it is possible to generate a map of the geographical distribution of patents, as well as the areas of the economy covered by possible applications of the research project. This, in turn, makes it possible to determine the geography and subject area of scientific and technical cooperation.

Conclusion

The tasks of the system for analyzing the relevance of scientific and technical projects developed in Techinformi are similar to the tasks of the Derwent Data Analyzer toolkit developed by Derwent for a wide range of patent research. The latest version of the system was put into operation in September 2021, and uses the Derwent World Patent Index (DWPI) extended bibliographic database as a database. The database contains more than 100 million patent documents and additional analysis and decision-making capabilities. The acquisition of this system would be extremely useful for the analysis and management of scientific and technical activities in Georgia.

ინფორმაციული მეცნიერება

საქართველოს საინფორმაციო სისტემის სამეცნიერო და ტექნიკური პროექტების შეფასება

ლ. ჩობანიანი*, ნ. მახვილაძე**, ფ. წოწკოლაური*, მ. ლებედევა*

* საქართველოს ტექნიკური უნივერსიტეტი, ინსტიტუტი ტექნიკური მეცნიერების, საინფორმაციო ტექნოლოგიების განვითარების განყოფილება, თბილისი, საქართველო

** საქართველოს ტექნიკური უნივერსიტეტი, ინსტიტუტი ტექნიკური მეცნიერების, ანალიზისა და საინფორმაციო რესურსების განყოფილება, თბილისი, საქართველო

(წარმოდგენილია აკადემიის წევრის რ. ხუროძის მიერ)

სამეცნიერო და ტექნიკური პროექტების შეფასებისა და ანალიზისთვის აუცილებელია საინფორმაციო სისტემების შექმნა, რომლის მეშვეობით შესაძლებელია თემატური კავშირის დადგენა მეცნიერებასა და ტექნოლოგიას შორის. ინსტიტუტი ტექნიკური მეცნიერების ადნიშნული საინფორმაციო სისტემა ფორმირდება სამეცნიერო და ტექნიკური პროექტების მონაცემთა ბაზისა და საერთაშორისო საპატენტო მონაცემთა ბაზების ინფორმაციის საფუძველზე. სამეცნიერო და ტექნიკური პროექტების ანალიზისთვის გამოყენებულ იქნა სისტემა SCOPUS-ის SciVal – ონლაინ ინსტრუმენტი, რომელიც ბიბლიომეტრული მეთოდების გამოყენებით, კვლევის შედეგებში შედეგის შესაძლებლობას იძლევა. SciVal-ონლაინ ინსტრუმენტი შესაძლებლობას იძლევა საპატენტო დოკუმენტების ციტირების მონაცემების საფუძველზე განისაზღვროს კავშირი სამეცნიერო კვლევებსა და საპატენტო კვლევებს შორის, რაც დამახასიათებელია სამეცნიერო კვლევების შედეგების პრაქტიკული გამოყენებისთვის. ანალიზისთვის მნიშვნელოვანია საერთაშორისო საპატენტო კლასიფიკატორის (IPC) რელევანტური ინდექსების და ყველაზე მეტად ციტირებული საპატენტო დოკუმენტების განსაზღვრა, რაც კვლევის პერსპექტიულობის დადგენის საფუძველია. მაღალრეიტინგული პუბლიკაციები და გამოგონებათა საერთაშორისო კლასიფიკაციის ინდექსები განთავსდება ტექნიკური მეცნიერების სამეცნიერო-ტექნიკური პროექტების მონაცემთა ბაზებში, როგორც პროექტის ბიბლიოგრაფიული აღწერის ცალკე ველი. მონაცემთა ბაზები ხელმისაწვდომია როგორც პროექტის შემსრულებლებისთვის, ასევე მეცნიერთათვის. კვლევის შედეგების თანახმად, ინფორმაცია პროექტის პერსპექტიულობის შესახებ, მიეწოდება სამეცნიერო-ტექნიკური პროექტის შემსრულებლებს. ინფორმაცია სარეკომენდაციო ხასიათისაა. ტექნიკური მეცნიერების სამეცნიერო და ტექნიკური პროექტების ანალიზის სისტემა და ამოცანები ანალიზურია Derwent-ის ინსტრუმენტარის, რომელიც Derwent-ის მიერ შემუშავებულია საპატენტო კვლევების ფართო წრისთვის.

REFERENCES

1. Narin F., Noma E. (1985) Is technology becoming science? *Scientometrics*, **7**, 3-6: 369-381.
2. Narin F. and Olivastro D. (1998) Linkage between patents and papers. An interim EPO/US comparison. *Scientometrics*, **41**, 1-2: 51-59.
3. Zatsman I.M., Havanskov V.A., Shubnikov S.K. (2013) Method of bibliographic information extraction from full-text descriptions of inventions. *Informatics and its Applications*, **7**, 4: 52-65.
4. Chyi-Shiang Hoo (2020) Impacts of patent information on clustering in Derwent Innovation's ThemeScape map, *World Patent Information*, **63**, 102001 <https://doi.org/10.1016/j.wpi.2020.102001>.
5. Matthis de Saint-Georges, Bruno van Pottelsberghe de la Potterie (2013) A quality index for patent systems, *Research Policy*. **42**, Issue 3: 704-719. <https://doi.org/10.1016/j.respol.2012.09003>.
6. Daniel Marques Golodne, Alexandre Sousa da Silva (2022) Managing patent examiners' workload by using patent application volume data, *World Patent Information*, **68**, 102095. <https://doi.org/10.1016/j.wpi.2022.102095>.

Received March, 2024