MiraBR – A System for Patent Analysis

Diogo Nolasco
Jonice Oliveira
diogo.sousa@ppgi.ufrj.br
jonice@ic.ufrj.br
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

Context: Innovation is essential for economic and technological development, with patents serving as a key metric to monitor and stimulate this innovation. However, patent information is often fragmented and distributed across isolated systems, hindering strategic and global analyses. Problem: The lack of integration between patent databases limits the ability to conduct comparative and evolutionary analyses, thereby restricting efforts to identify trends, prospect opportunities, and monitor international competitiveness. Solution: The MiraBR system was developed as an open-source web application that collects, integrates, and analyzes patent data from multiple national and international sources. Using a web crawler, data is extracted from databases such as INPI, Derwent, and Espacenet, processed, and integrated into a centralized database. The system supports multi-language search queries and enables analyses based on classifications, countries, inventors, and patent holders. It also provides intuitive visualizations that facilitate strategic and comparative analysis. IS Theory: This work is conceived under the Socio-technical theory, as the system integrates human and technical elements. Method: The research involved a proof of concept to validate the system's core functionalities and a quasi-experiment to assess its usability and effectiveness in real-world scenarios. The quasi-experiment included specialists from the intellectual property field, who evaluated the system's performance and relevance through practical application and structured questionnaires. The analysis approach was descriptive, aiming to present the system's capabilities and its alignment with user needs, with a focus on quantitative metrics (e.g., precision and recall) and qualitative feedback (e.g., usability assessments). This mixed-method approach ensured a comprehensive evaluation of the system, addressing both technical performance and user satisfaction. Results: The system was evaluated by INPI experts and received high approval for its ease of use, rich visualizations, and quality of the analyses. Precision and recall metrics indicated satisfactory performance in information retrieval. Contributions: MiraBR contributes to innovation management by enabling integrated strategic analyses of patents, enhancing technological support for science and innovation in Brazil. It also paves the way for future research in data mining and visualization applied to the domain.

CCS Concepts

• **Information systems** → *Information systems applications*; Information integration; • **Social and professional topics** → **Patents**; • **Human-centered computing** → **Information visualization**.

Keywords

Information Visualization, Patents, Data Management

1 Introduction

Information on the knowledge produced by countries is maintained by agencies and organizations, with patents serving as one of the most significant indicators of innovation. In this context, one of the most important items is the registration of patents, which is a public concession granted by the state that guarantees its holder the exclusive right to commercially exploit their creation, becoming a valuable measure of innovation generated by a country [22]. The data made available by the main patent databases are presented to the user in their almost raw state, which can be accessed through simple queries and without any integration mechanism. Brazil is at an average stage of scientific and technological development and has a high demand to transform its scientific production into applied innovation [3, 6].

The study of patents provides innovation metrics and acts as an indicator of a country's economic development [23]. [14] define the following subareas of patent research: analysis of emerging technologies and technology dynamics (trend analysis); technology forecasting; roadmapping and forecasting; research and development management; indicators of engineering industries; science and technology (S&T); evolutionary economics; technology assessment and impact; science policy analysis and studies; and knowledge flow analysis.

The use of technology as a form of support of such researches can enable improvements in the way data is stored, maintained, and consulted. The analysis and visualization of innovation data, both in the local and international context, could be of great value in identifying demands, outlining development plans, assessing competition, and forming partnerships for technological development. For this to be possible, an integrated view of the entire innovation production ecosystem and new means of supporting analysis in science and technology are required.

Due to the problems of analyzing patent data and the strategic importance of scientific and technological management for the country's development, we present the proposal entitled MiraBR. This solution aims to assist in the management of knowledge produced by inventors or organizations, allowing the import and aggregation of patent data, in addition to more refined analyses and visualizations. It is based on the Socio-technical IS Theory as a system that integrates human and technical elements. It enables decision-makers (social) to interact with and benefit from a technically integrated patent data system. It offers socio-technical solutions by merging

technical integration (data aggregation and visualization) with the social need for strategic patent management.

Previous works have shown some approaches for the area of innovation with the use of patent data, such as in [24], where the authors aim to generate queries for patent data retrieval and [20], where the authors build a recommendation system for similar patents. Others such as [13, 19] work on the analysis by mapping the technological area for a domain or by methods for technology forecast. All of them are limited by the quality and completeness of the data used, specifically for Brazil, where patent documents cannot be retrieved many times due to outages or data format changes.

The rest of this article is organized as follows. Section 2 discusses the work related to that presented in this article. Section 3 presents the theory for contextualizing the system's theme and domain. Section 4 presents the proposal and explanation of the system's aspects. Section 5 presents the results and the evaluation of the proposal. Finally, Section 6 concludes with final considerations.

2 Related Works

The patent domain have studies that range from retrieval and methods to analysis and integration. In [16] the authors do a comparative analysis of the databases and search engines available for data retrieval. They compare proprietary software with online databases such as Espacenet and Patentscope to assess which have more features and return more documents when searching. MiraBR aims to aggregate data from online databases and unlike proprietary systems the system is extensible by adding more data. [11] also compares the online free databases and specifically the data concerning the INPI (Instituto Nacional da Propriedade Industrial) database, showing that the brazilian patents are more available on this base.

An ontology and linked data approach for patents are presented by [17]. They extend the USPatents ontology and use linked data to build a citation map based on data such as inventor, applicant and the explicit citations between documents. Similarly, [18] creates a semantic network based on US patent data, using natural language processing to extract terms from text. Although both have good results, they present a manual intensive process limited by the format of the data used. Our proposal provides visualizations on the data across various data dimensions.

Finally, recent analysis works such as [19] provide a systematic mapping of technology production in the COVID-19 domain, comparing national technological production with the international scenario using patent documents from 2020. Analysis like this could be more easily produced by MiraBR which aggregates data from national and international databases and is multilingual in its search. Another common use for patent data is technological prospecting and technology forecasting. In [13], the authors propose a method for forecasting using link prediction to identify potential trends based on the links between documents. Using data from various sources could also enhance the prediction and decision-making, specially in a global scenario.

3 Theoretical Background

Creating a new product is generally an expensive process. Protecting the new product so that competitors do not copy the invention is the purpose of applying for an invention patent [15]. A patent is "a formal document, issued by a government agency, through which property rights and exclusive use of an invention described therein are granted and recognized. It is a privilege granted by the state to inventors (individuals or legal entities) or holders of the right to invent or improve products and manufacturing processes" [5].

To obtain the rights to an invention, it is necessary to demonstrate to the state the intention to obtain it, through an invention registration process, which may vary in its properties and particularities according to the country where the legal procedure for the application is carried out [10].

The definition of what is an invention is vague and may include a variety of objects and ideas. However, legally, in order to be patented, it must comply with three mandatory requirements: novelty, inventive activity and industrial application [9].

3.1 Patent classification

The International Patent Classification [21] categorizes the invention into one of 8 (eight) main sectors, which are:

- Section A - Human Necessities (Vol. 1)
- Section B - Processing Operations; Transportation (Vol. 2)
- Section C - Chemistry and Metallurgy (Vol. 3)
- Section D - Textiles and Paper (Vol. 4)
- Section E - Fixed Constructions (Vol. 5)
- Section F - Mechanical Engineering / Lighting / Heating / Weapons / Blasting (Vol. 6)
- Section G - Physics (Vol. 7)
- Section H - Electricity (Vol. 8)

The complete classification symbol consists of symbols representing Section (as above), Class (number composed of two digits), Subclass (lower case letter), Group and Subgroup. The class, subclass, group and subgroup represent hierarchical levels within a section, as shown in the example below:

Ex:

Classification - A01B 1/24

- A - Human Needs Section
- 01 - Class - Agriculture; Forestry; Livestock; Hunting; Trapping; Fishing
- B – Subclass - Soil work in agriculture or forestry; Parts; Details or accessories of agricultural machines or implements, in general
- 1/00 - Main Group - Hand implements
- 1/24 – Subgroup - Combinations of different species of hand implements for treating meadows or lawns

3.2 Patent databases

This work used and integrated information from three patent databases: INPI, Derwent, and Espacenet.

In the INPI database [4], the information provided is the registration number, title, abstract, status, country, filing and publication dates, inventors, owners and, in the case of documents published

since 2006, it is possible to access the original document in PDF format. All information contained in the registry is described in Portuguese.

The Derwent World Patent Index [1] is a product of Thomson Reuters and is made available in Brazil by CAPES. It covers more than 14.3 million basic inventions from 40 patent authorities around the world. In addition to the international classification, it has its own classification. Its search tool is capable of using a set of logical operations and recording searches and their results for later use.

Espacenet [2] is a product of the European Patent Office, through which it is possible to search its EPODOC database and access original patent documents. The patent database collection is available free of charge on the internet and consists of approximately 60 million patent documents from more than 80 countries, including patent applications filed and published in Brazil.

The INPI database was chosen because of its relevance to the national scenario, and the Derwent and Espacenet databases due to their importance and because they contain information from Europe and the majority of countries in the world.

Although all of these databases represent information from patent documents, the data made available varies from one to another. In the INPI database, we have the information in Portuguese, while in the others, it is only possible to retrieve data in English. In Espacenet, we have the European classification, and in Derwent, its own classification, a registration number and abstracts that are rewritten. In addition, there are several differences in formatting, for example, the registration status or filing date accompanying the patent number; inventors listed or in bibliographic format; classification divided into types (International, European). Each of these databases has a combination of these particularities, which makes integration between them a challenge, as well as the extraction of their information itself.

3.3 Socio-Technical Systems Theory and Its Application

The Socio-Technical Systems Theory (STS) emerged as a framework to analyze and design systems that integrate both technical and social dimensions. It emphasizes the interdependence between human actors, organizational processes, and technological infrastructures, advocating for a holistic approach to system development and evaluation. STS theory argues that the success of an information system depends not only on its technical performance but also on how well it aligns with the social context and organizational dynamics in which it operates.

Key principles of STS theory include [12]:

- Joint optimization: Balancing social and technical components to achieve system effectiveness and user satisfaction.
- User-centered design: Ensuring systems are developed with a deep understanding of user needs and contexts.
- Adaptability: Designing systems that can evolve in response to changing social and technical environments.

In the context of this study, MiraBR exemplifies the application of STS principles by addressing both the technical challenges of patent data integration and the social needs of its users, such as researchers, policymakers, and organizations. The system:

- Integrates heterogeneous data sources: Resolving technical issues of interoperability across patent databases like INPI, Espacenet, and Derwent.
- Facilitates user engagement: Providing intuitive visualizations and multilingual search capabilities tailored to the needs of diverse stakeholders.
- Supports strategic decision-making: Enabling users to derive actionable insights from complex data, aligning technical functionalities with social and organizational objectives.

Moreover, MiraBR acknowledges the dynamic nature of the innovation ecosystem. By incorporating feedback from intellectual property specialists during its evaluation phase, the system adheres to STS principles of adaptability and user-centered design. This iterative approach ensures that the system evolves in response to user demands and technological advancements.

Through the lens of socio-technical theory, MiraBR serves as a practical example of how information systems can bridge the gap between technical innovation and social utility. By optimizing the interplay between its technical architecture and the needs of its users, the system contributes to the broader goals of fostering innovation and enhancing intellectual property management in Brazil and beyond.

4 Proposal

MiraBR is an open source web application that allows users to import and aggregate data from external databases (national and international), analyze data in its raw form, and use information visualization models to identify relationships between research studies, communicating information effectively and intuitively. A diagram of the entire process - from data collection to its integration, which is necessary for data analysis and visualization - can be seen in 1. In this section, we will explain the main parts of the system.

4.1 Import Process Flow

The first step in using the solution is to extract data from each database to be used. To do this, we need to define a subject of interest, a domain, and from there define terms to be searched. Once this is done, the search for the terms is performed in each of the databases.

The searches are performed using a web crawler (a programmed robot that collects information on the internet), which extracts information from patent search portals by evaluating the HTML structure of the pages. The information that will make up the integrated database is extracted and the results are saved in temporary files in csv format. It is important to remember that this data, despite having already undergone an initial treatment for csv formatting, is still different from each other depending on the source databases.

Next, the data import process takes place, which is responsible for reading the files in csv format, performing the other necessary treatments, integrating them and storing them in the database, the diagram of which is shown in 2.

Through scheduled tasks, the data import process reads the files, performs a first cleaning stage on the extracted text, encoding special characters that may cause errors in the database when inserting the data. In the second stage, the dates contained in the documents are processed (making the formats compatible) and performing the

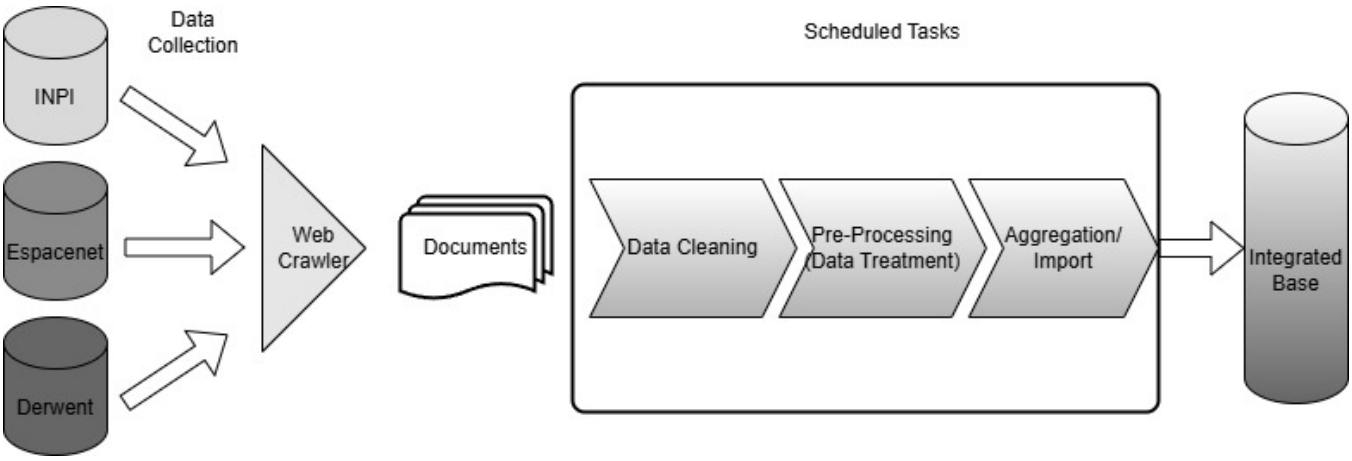

Figure 1: Data Fusion process flow

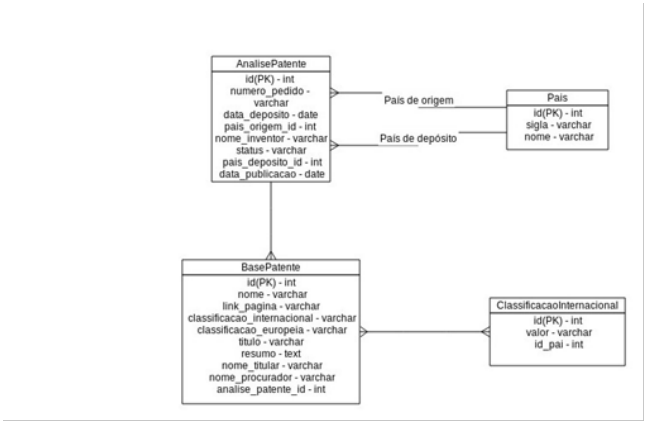

Figure 2: Database Schema

separation of information, for example, "BR20100303A" becomes number "BR20100303", country of filing "BR" and status "A". This and other treatments are performed using Natural Language Processing via regular expressions, which separate the texts using syntax rules.

After all these steps, the information for each patent is saved as new records in the integrated database, with homogeneous fields as shown in Figure 2.

4.2 Search and Analysis

In MiraBR, it is possible to perform aggregated queries based on the following attributes: Search (the term to be searched), Classification, Country of origin, Country of filing, Status, Chart (the type of), Filing date, Publication date, Inventors and Owners.

In the search, the specified term is searched in the title and abstract of the documents. The search can be done using terms written in either Portuguese or English. The translation (Portuguese <> English) of the term is done through the MyMemory[1] web service. It receives the terms and returns the possible translations in order of probability (the first in the order is chosen as the translation). Both terms (original and translated) are used in the search. For example: The term "Stem cell" is translated to "Célula tronco" and the set formed by these two terms will be searched. Thus, in a transparent manner to the user, the search is expanded so that documents from the three databases are retrieved.

The classification can be searched by section, class and subclass, which are related to the divisions illustrated in the previous section.

The country of origin includes the place of origin of the invention. An invention can also be patented in other countries, so the country of filing is registered to cover the places where patents are filed. Both locations (origin and filing) can be searched.

By status, as the name suggests, allows searching by the state of the patent, for example, whether its validity period has expired or not. A complete list of possible states can be found in [22].

The default graph is the geographic one, which shows the distribution of the patents searched by country. The user can choose from other types of visualization (pie charts, area, bar, column, line or table).

The filing date is the date on which the patent application was filed. After the entire patent application process, it is published on a publication date. Both dates can be searched together with the other fields.

The search by inventor locates inventions by the names of the creators. They do not always hold the rights to their inventions, which is up to the owner, who can also be searched.

Through these fields, we can have broader or more refined searches and, in addition to viewing graphs, we can switch to a list view of the results. Figure 3 shows an example of the search result for the term "stem cell" with a view of the list of patents found in the background. When we select a document from the list, a page is displayed with detailed information about the patent, which can be seen in Figure 3. Note that in the details, the information is separated by database, and it is even possible to access the original document in the database portals.

Figures 4 to 6 show the results of the search for the term "stem cell" using the graphical view. Figure 4 shows the results in a tabular

[1] https://mymemory.translated.net/

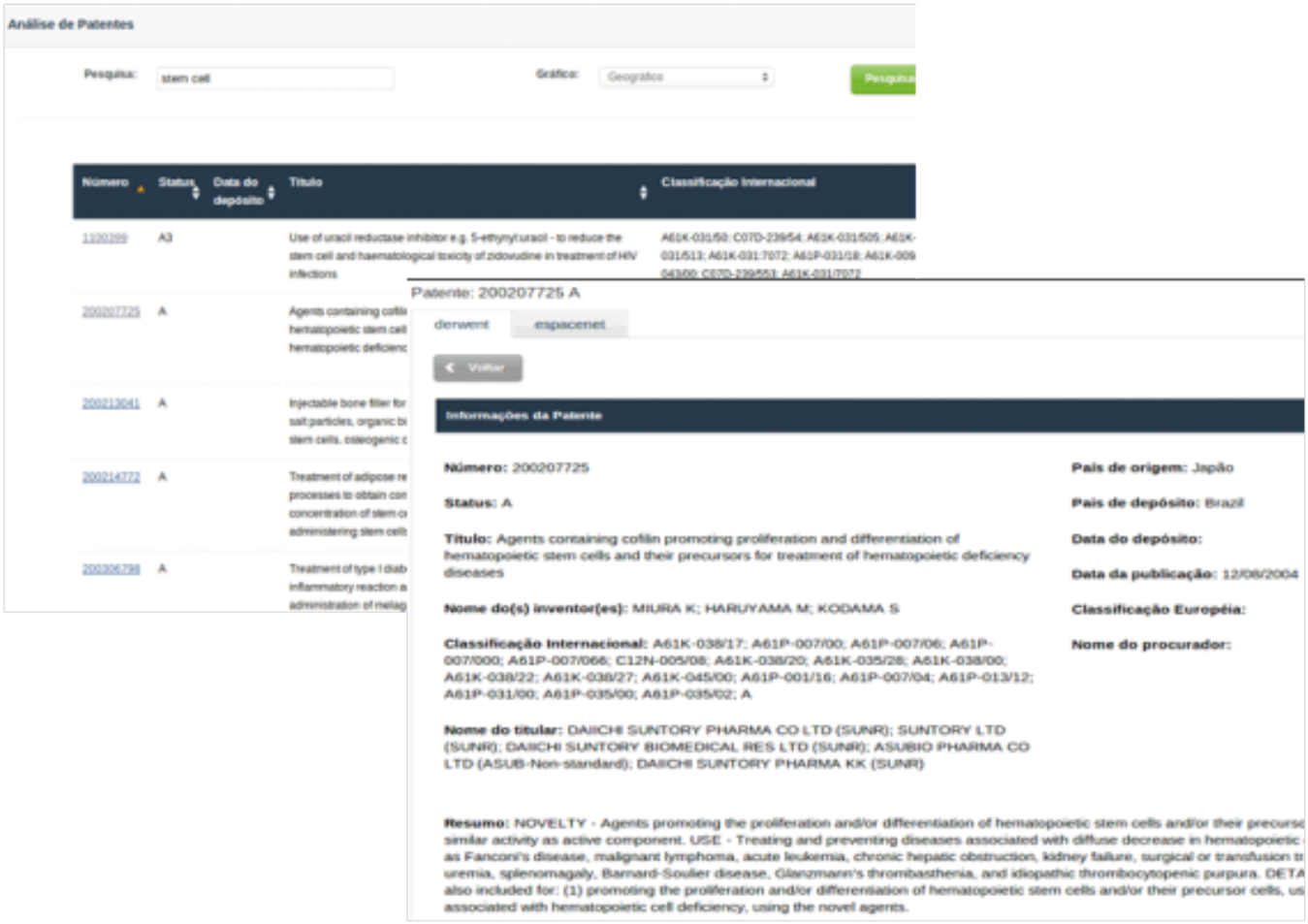

Figure 3: Search Result (back) and Patent Details (front)

view in the background and in a pie chart. Figure 5 shows the results in the standard graph (Geographic), with the distribution of the patents by country. In figure 6 we have a comparison of the countries' production in a column chart view.

5 Evaluation

To evaluate the proposal, a quasi-experiment and a quantitative evaluation were carried out through precision and recall calculations.

5.1 Quasi-experiment

This assessment was aimed to analyze, from the perspective of the end user, the applicability and importance of the solution. The participants were two INPI employees who had worked as industrial property researchers for seven years and dealt directly with patent analysis in their daily activities, although they were from different sectors. A sample of 20 terms was used to create the integrated database, which were suggested by the participants themselves according to their area of expertise.

The terms used were: "stem cell", "stemcell", "progenitor cell", "embryonic cell", "embryogenic cell", "wind generation", "wind farm", "wind turbine", "wind power", "wind energy", "eolic generation", "eolic farm", "eolic turbine", "eolic power", "eolic energy", "windmill generation", "windmill farm", "windmill turbine", "windmill power" and "windmill energy". After a quick training session, the participants used the system and, after using it, evaluated the solution through a questionnaire.

According to the analysis of the response data, it was found that the respondents classified the solution as very easy to use, due to the fact that the interface resembled already known search systems. The search options were considered very satisfactory by one, while the other demonstrated that there was still the possibility of approaching many other forms of research. The result obtained through the search mechanisms developed was considered very satisfactory by both participants. The geographic visualization and the number of visualization options were very satisfactory for the respondents.

The aggregation of information was a highly appreciated item and the use of visualizations facilitated the global perception of the innovations produced in each country according to the participants'

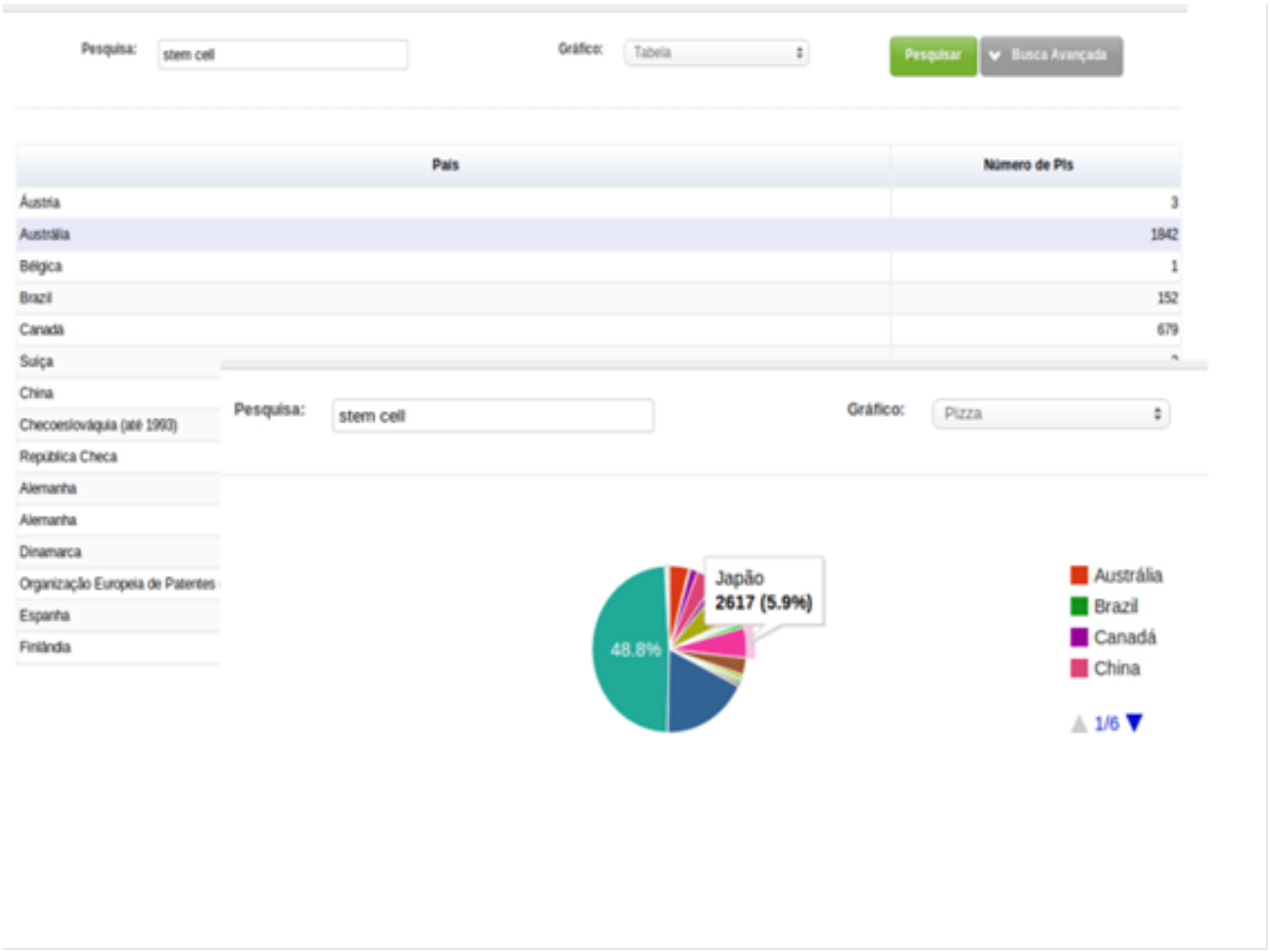

Figure 4: "Stem cell" patents by country - table visualization (back), and by country percentage - pie visualization (front)

opinion. Regarding the usefulness of the graphs, we saw that the geographic, bar, pie and column graphs were considered the most useful for this type of analysis.

5.2 Quantitative Evaluation

The second assessment compared the results obtained by MiraaBR (using the integrated database from the previous assessment) with two reference databases, which were manually constructed by INPI professionals themselves. To create the reference databases, the experts searched for, analyzed and classified the records containing the terms "stem cell" and "wind power" in the Derwent and Espacenet databases. These reference databases contained the relevant documents, according to each researcher, according to whether or not the subject of the document was included in the area being researched. For example, a document containing the term "stem cell" is considered relevant when its subject is associated with the knowledge area of stem cells.

To assess the quality of the search system, the following information retrieval metrics were calculated: Precision and recall [7]. These are given by the formulas:

$$Precision = \frac{\text{Total relevant records retrieved}}{\text{Total records retrieved}}$$

$$Recall = \frac{\text{Total relevant records retrieved}}{\text{Total relevant records}}$$

Using the Derwent and Espacenet databases as the data source, 524 patent applications related to the topic "stem cell" were found, of which 178 were relevant. After performing a search on MiraBR, using the same term, 411 documents were retrieved, of which 156 were relevant. With this, we had: Precision = 156/411 = 0.3795; Recall = 156/178 = 0.8764.

Although the precision is relatively low, it did not differ much from the precision obtained through the official databases.

For the term "wind power", 11 documents were retrieved, of which 9 were relevant out of a total of 30, resulting in Precision

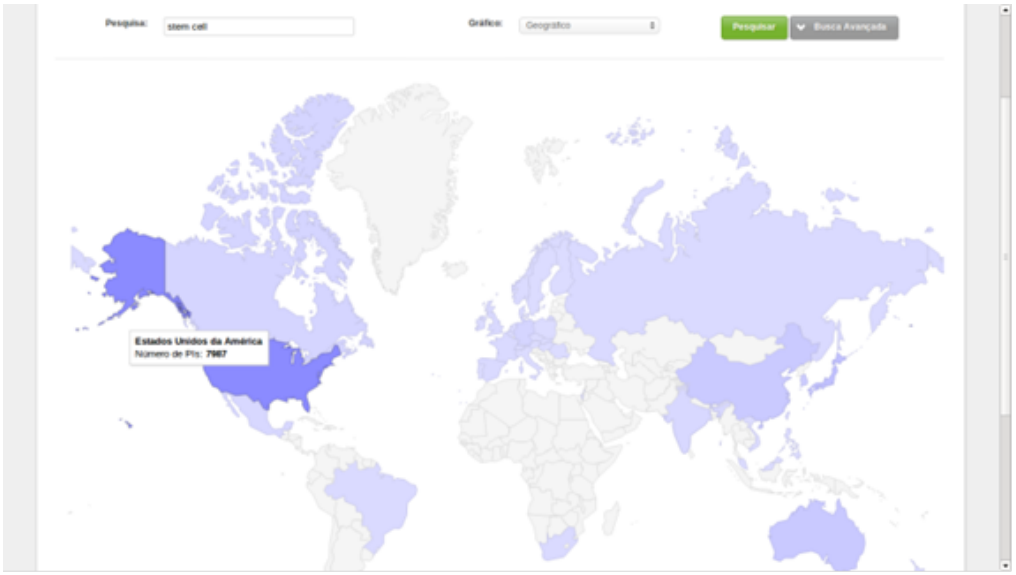

Figure 5: Geographic Visualization of "stem cell" patents by quantity

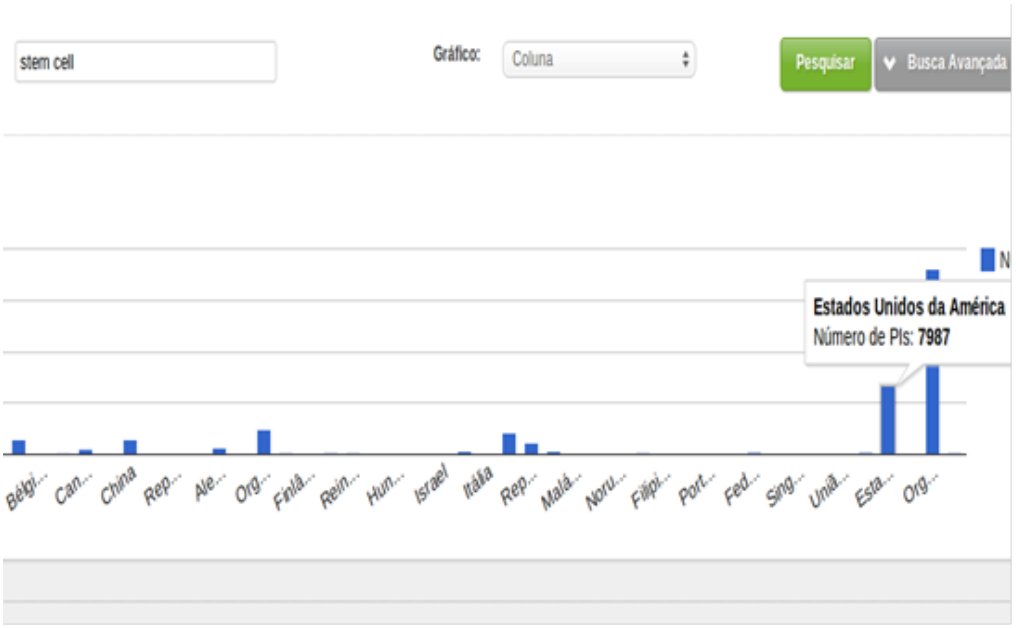

Figure 6: Comparison of the applicant countries on "stem cell" patents - column visualization

= 0.8181 and Recall = 0.3. In this case there was an inversion in values, which can perhaps be explained by the greater scope of the term “wind Power” and the fact that many documents related to wind energy do not explicitly contain the term in question.

6 Conclusion

This work presented an alternative proposal for analyzing intellectual property, capable of integrating data from several databases and generating visualizations that show the patent situation around the world in a simple and efficient way. It is capable of performing richer queries through compound searches and facilitating the understanding of results through visualizations. Therefore, it is important to provide the necessary technological support for scientific development and national innovation.

Evaluations were carried out through a quasi-experiment and the calculation of precision and recall information retrieval measures. These showed that the tool is useful in the field of patent analysis and that it is possible to create useful integrated and multilingual

databases. This form of analysis based on integrated data broadens the horizon for the creation of scientific cooperation between research areas and countries, in addition to improving the global vision of innovation production from different perspectives.

The proposal makes significant contributions to information systems applied to the domain and ecosystem of science, technology, and innovation by addressing a critical gap in the integration and analysis of patent data. The system not only enables a global and strategic view of technological production but also supports the identification of trends, market opportunities, and technological prospecting. Additionally, it broadens analytical capabilities for various stakeholders, including researchers, organizations, and policymakers. This work demonstrates that integrated analytical systems can be both accessible and effective, even in resource-constrained environments. Additionally, MiraBR aligns with the Grand Challenges in Information Systems [8], specifically with the challenge of Systems of Information Systems by creating a unified platform that combines multiple patent databases into a cohesive system. This approach exemplifies how interconnected systems can enhance functionality and support strategic decision-making, making it a practical implementation of the "systems-of-systems" paradigm in the context of innovation management.

Limitations of the proposal include its dependency on external databases and the challenges associated with maintaining interoperability as these systems evolve. Additionally, while the system currently supports multilingual queries and visualizations, further refinements in user interface design and customization are needed to maximize usability.

The approach opens up many possibilities for expanding future research, such as identifying better forms of visualization according to the type of research and improving the identification of areas through text mining techniques. Improvements such as incorporating artificial intelligence and machine learning to enhance data classification and trend prediction. Finally, efforts that aims to integrate additional data sources and exploring new metrics for evaluating innovation impact.

Acknowledgments

To Dr. Priscila Rohem (INPI) for providing the data and the other members of INPI who participated in the evaluation.

References

[1] [n. d.]. Derwent World Patents Index (DWPI) | Clarivate — clarivate.com. https://clarivate.com/intellectual-property/patent-intelligence/derwent-world-patents-index/. [Accessed 18-11-2024].
[2] [n. d.]. Espacenet. https://worldwide.espacenet.com/patent/. [Accessed 18-11-2024].
[3] [n. d.]. Foundations for rebuilding Brazil's scientific capacity — revistapesquisa.fapesp.br. https://revistapesquisa.fapesp.br/en/foundations-for-rebuilding-brazils-scientific-capacity/. [Accessed 18-11-2024].
[4] [n. d.]. INPI - Instituto Nacional da Propriedade Industrial. https://www.gov.br/inpi/pt-br. [Accessed 18-11-2024].
[5] [n. d.]. Patentes. https://www.gov.br/inpi/pt-br/acesso-a-informacao/perguntas-frequentes/patentes. [Accessed 18-11-2024].
[6] [n. d.]. Science, Technology, and Innovation in Brazil — unesco.org. https://www.unesco.org/en/node/104636. [Accessed 18-11-2024].
[7] Ricardo Baeza-Yates, Berthier Ribeiro-Neto, et al. 1999. *Modern information retrieval*. Vol. 463. ACM press New York.
[8] Clodis Boscarioli, Renata Mendes de Araujo, Rita Suzana Maciel, Valdemar Vicente Graciano Neto, Flavio Oquendo, Elisa Yumi Nakagawa, Flavia Cristina Berrnardini, José Viterbo, Dalessandro Vianna, Carlos Bazilio Martins, et al. 2017. I GranDSI-BR: Grand Research Challenges in Information Systems in Brazil 2016-2026. (2017).
[9] BRASIL. 1996. Lei 9279, de 14 de maio de 1996. Publicada no Diário Oficial da União em 15 de maio de 1996. [Accessed 18-11-2024].
[10] BRASIL. 1998. Lei 9609, de 19 de fevereiro de 1998. Publicada no Diário Oficial da União em 20 de fevereiro de 1998. [Accessed 18-11-2024].
[11] Paula Cotrim de Abrantes, Daniel Barros Junior, Elisângela Aparecida da Silva, Genízia Islabão Islabão, and Douglas Alves Santos. 2022. Estudo comparativo de bases gratuitas de patentes: patentscope (wipo), espacenet (epo), buscaweb (inpi/br). *Revista Tecnologia e Sociedade* 18, 54 (2022), 125–142.
[12] Fred E Emery and Eric L Trist. 1960. Socio-technical systems. *Management science, models and techniques* 2 (1960), 83–97.
[13] Patrick Kochan, Bartholomeo Barcelos, and Alexandre Gonçalves. 2024. Identifying Technological Trends: A Patent Analysis Method for Technology Forecasting. In *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados* (Florianópolis/SC). SBC, Porto Alegre, RS, Brasil, 327–340. https://doi.org/10.5753/sbbd.2024.240268
[14] Farshad Madani. 2015. 'Technology Mining'bibliometrics analysis: applying network analysis and cluster analysis. *Scientometrics* 105, 1 (2015), 323–335.
[15] Suzana Pinheiro Machado Mueller and Valmira Perucchi. 2014. Universidades e a produção de patentes: tópicos de interesse para o estudioso da informação tecnológica. *Perspectivas em Ciência da Informação* 19 (2014), 15–36.
[16] Edilson Araújo Pires, Nubia Moura Ribeiro, and Cristina M Quintella. 2020. Sistemas de busca de patentes: análise comparativa entre espacenet, patentscope, google patents, lens, derwent innovation index e orbit intelligence. *Cadernos de prospecção* 13, 1 (2020), 13–13.
[17] Lucas Resende, José Pérez-Alcázar, and Esteban Tuesta. 2021. Desenvolvimento de uma Ontologia de Patentes e criação de uma base de dados ligados para patentes brasileiras.
[18] Serhad Sarica, Jianxi Luo, and Kristin L Wood. 2020. TechNet: Technology semantic network based on patent data. *Expert Systems with Applications* 142 (2020), 112995.
[19] Sandro Luis Freire de Castro Silva, Maria Augusta Silveira Netto Nunes, Marcelo Fornazin, and Rodrigo Pereira dos Santos. 2021. The profusion of information systems to combat the covid-19 pandemic: A systematic mapping of the state of the art and brazilian challenges of technological production. In *Proceedings of the XVII Brazilian Symposium on Information Systems*. 1–8.
[20] Amy Trappey, Charles V Trappey, and Alex Hsieh. 2021. An intelligent patent recommender adopting machine learning approach for natural language processing: A case study for smart machinery technology mining. *Technological Forecasting and Social Change* 164 (2021), 120511.
[21] WIPO. 2022. International Patent Classification (IPC) — wipo.int. https://www.wipo.int/publications/en/details.jsp?id=4582. [Accessed 18-11-2024].
[22] WIPO WIPO. 2023. Intellectual Property Handbook: Policy, Law and Use, WIPO Publication no. 489 (E)(2008).
[23] Yen-Chun Jim Wu and Hsing-Piao Liu. 2006. Technological innovation assessment of business-to-business electronic marketplaces. *Journal of the American Society for Information Science and Technology* 57, 8 (2006), 1093–1104.
[24] Xiaobing Xue and W Bruce Croft. 2009. Automatic query generation for patent search. In *Proceedings of the 18th ACM conference on Information and knowledge management*. 2037–2040.