Analysis of the Information Recorded in Patents using Bibliometrix

Eduardo Amadeu Dutra Moresi[*]

Department of Governance, Technology, and Innovation, Catholic University of Brasilia, Brasilia, Brazil

INTRODUCTION

This paper presents an advanced methodology for exploring patent application documents using the R-Bibliometrix package, an instrument used for bibliometric analysis. It delves into the process of meticulously extracting valuable information from patent documents for qualitative analysis. These documents are a treasure trove of knowledge about emerging technologies and research trends. The study ambitiously addresses the research question: How can bibliometric analysis are employed to interpret documents obtained from patent databases? This question is tackled through an analysis of patents extracted from the Lens database, a globally accessible, open database housing over 130 million patents from over 100 jurisdictions worldwide.

The research outlines a systematic analysis approach that includes steps like formulating the research design, collating and compiling patent data, preparing and importing metadata, performing bibliometric analysis (encompassing both performance analysis and technology mapping), conducting qualitative analysis, and finally, reporting and disseminating research findings. The paper underscores the importance of patent classification codes, such as the International Patent Classification (IPC) and the Cooperative Patent Classification (CPC). These codes are instrumental in dissecting the technology life cycle and identifying trends in technological evolution. However, it highlights a limitation of the R-Bibliometrix package, its inability to analyze patent data due to disparities in metadata tags, and proposes adapting these tags to match those commonly used in academic work.

The paper skillfully demonstrates the patent analysis method by applying it to the realm of virtual worlds. A virtual world is defined as a simulated environment where agents can interact in a persistent, uninterrupted virtual setting. The article highlights the plethora of terms used interchangeably to describe similar technologies in literature, leading to a fragmented understanding of virtual worlds.

DESCRIPTION

For this study, the Lens database was used to search for patents related to virtual worlds that were submitted and granted. The comprehensive search yielded 6565 documents, which included 1952 granted patents, 76 limited patents, and 4537 patent applications [1]. The study discloses that the peak of patent production for virtual worlds occurred in 2020, with an impressive 490 documents published and an annual growth rate of an 18.96%. The United States, China, and the World Intellectual Property Organization (WIPO) were the primary jurisdictions for patent registration. IBM, Microsoft, and searete were the leading patent owners, with searete having the highest number of affiliated inventors, followed by Samsung and IBM [2].

In this investigation, the CPC codes found in numerous documents were identified and studied using exploratory thematic analysis and thematic mapping techniques. The research unveiled 4021 CPC subgroup codes, 634 group codes, 169 subclass codes, and 59 class codes. The most frequent CPC class codes were "G06 computing calculating or counting" and "A63-sports; games; amusements".

Thematic mapping was used to identify technology trends and conceptual structures, incorporating Callon's measures of centrality and density to map research themes on a two-dimensional strategic diagram. This diagram is divided into four distinct quadrants: motor themes, basic themes, emerging or declining themes, and niche or specialized themes. The thematic map of the CPC subclass codes unveiled certain codes as basic, emerging, niche, and motor themes [3]. The bigrams extracted from patent abstracts also showed that terms related to computing were categorized as motor themes, while terms like virtual world were recognized as basic themes.

The R-Bibliometrix package has a unique feature called thematic evolution, which aids in interpreting the conceptual structure of technologies over time. The paper uses diagrams to depict the evolution of the technical structure of CPC code subclasses over five periods from 1990 to 2022. It also charts the thematic evolution of technology structure based on abstract bigrams.

This theme is consistent throughout, and the study of the flow between periods provides an understanding of the conceptual trajectories of technology during each discrete time slice [4].

The paper suggests inclusion criteria for selecting a patent set for comprehensive content analysis. A sample of documents must be systematically read and scrutinized using content analysis techniques based on specific questions. The qualitative analysis allows the definition of dimensions, categories, and subcategories of analysis, assisting in the creation of conceptual models and answering research questions.

CONCLUSION

The study concludes by highlighting the potential of the R-Bibliometrix package for in-depth patent analysis and encourages further exploration of its capabilities. It recommends gaining a deeper understanding of the R-Bibliometrix package, particularly in terms of interpreting the tags exported by the Lens database for a patent and their correspondence with academic work tags. This understanding will facilitate further analysis of the qualitative data present in the patents. The work could be enriched if experts in the examined field jointly conduct a qualitative analysis to address complex questions that may arise from the discussion of this exploratory study.

REFERENCES

1. Moresi EAD, Pinho I. Analysis of the conceptual structure of information recorded in patents. World Conf Qual Res. 2023;688:209-228.

2. Jin L, Sun X, Ren H, Huang H. Hotspots and trends of biological water treatment based on bibliometric review and patents analysis. J Environ Sci. 2023;125:774-785.

3. Gupta VK, Pangannaya NB. Carbon nanotubes: Bibliometric analysis of patents. World Pat Inf. 2000;22(3):185-189.

4. Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools. World Pat Inf. 2008;30(2):115-131.