Information Retrieval through Online Public Access Catalogue (OPAC): A Bibliometric Analysis

Hamza Zerrouak[1], Sarra titila[2]

[1]Phd Student, University of Mohamed Khider Biskra, College of Humanities and Social Sciences, Department of Humanities, Division of Library Science, Specialization Technology and Information Engineering, Social Change and Public Relations in Algeria laboratory (Algeria).

[2]Senior Lecturer A, University of Mohamed Khider Biskra, Specialisation in Information Technology in Document Systems, New Information Technologies and their Role in National Development Laboratory, Constantine2 (Algeria).

The Author's Email: Hamza.zerouak@univ-biskra.dz[1], Sarah.titila@univ-biskra.dz[2]

Received: 05/2023; Published: 01/2024

Abstract:

Information and technology are the driving force of progress and the fuel of the era, prompting libraries to keep pace with various technological advancements in order to provide access and retrieval of information sources and bibliographic data to their beneficiaries to meet their knowledge and information needs. This leads us to discuss information retrieval systems through online public access catalogues (OPACs).

Through a bibliometric analysis, this study analysed all documents from 2000 to 2022 within the Scopus database, using the Vosviewer software for a total of 95 publications. The study identified the journals, countries and authors with the most interest in the field, as well as the keywords and their relationships. The results of the study revealed the fluctuations and lack of interest in the field, as well as its narrow specialisations and precise terminology. The study also highlighted the high interest of the United States, its journals and even its researchers in the field, which surpasses and varies greatly compared to other countries represented, such as South Africa, Greece, Singapore and India.

Keywords: Information Retrieval, Online Public Access Catalogue (OPAC), Bibliometric Analysis.

Introduction:

Societies have moved from the Industrial Revolution to the Information Revolution, in which the Internet has been a qualitative leap and a stage for scientific and information competition, forming a new territory. According to Schwartz (1999), the Internet is considered as a new culturally autonomous country or region[1]. However, it has given rise to various conflicting opinions,

[1]- Schwartz, D. (1999), "The Internet as an agent for social and organizational change", Internet Research, Vol. 9 No. 3. https://doi.org/10.1108/intr.1999.17209caa.001

particularly with regard to information retrieval. Massimo Melucci, Maristella Agosti and others affirm that the Web environment has evolved in a way that makes it difficult for anyone to try to explore all the information distributed on its pages. By mid-1999, there were 800 million pages containing 6 terabytes of textual data. This necessarily requires support and the availability of tools that can help end-users to select the most relevant web pages to answer specific information requests[2]. This leads us to talk about online library catalogues, which according to Fox (2007) are one of the most important platforms for reliable scientific information. Perhaps OPACs have been a source of inspiration for many advanced services found on the Internet today[3].

While libraries are trying to find their place in the digital space by providing information in bibliographic or full-text format through their online catalogues, most researchers rely on search engines and artificial intelligence techniques, which pose several challenges. One of the main challenges is the difficulty of performing a strong semantic analysis of large text collections or multimedia objects, given the unstructured nature of natural languages, which requires a "bag of words" approach and high user skill in formulating research queries, as highlighted by Thomas Mandl's study (2008)[4]. This leads us to investigate and understand the extent to which researchers are inclined to publish on information retrieval through these online library catalogues, known as Online Public Access Catalogues (OPACs), due to the valuable and rich information or data they provide to access valuable and valuable sources, thereby promoting scientific and academic research in the midst of technological and technical competition.

In this context, this paper aims to analyse the publications in the Scopus database over the last two decades in relation to information retrieval through OPACs, using bibliometric methodology. The data and citations of prestigious and scientific journals have been collected in order to understand the importance of this topic in the light of technological changes, new research areas and information retrieval in the digital environment. It also provides insights into the leading researchers, the relevant journals and the scientific links between them, by addressing the following questions:

1. What is the distribution of information retrieval through OPACs between the years 2020-2022?
2. Which journals and authors will be the most relevant for information retrieval via OPACs?
3. Which countries are the most productive in research related to information retrieval via OPACs?

[2]- Agosti, M., Crestani, F., & Pasi, G. (Eds.). (2001). Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000 Varenna, Italy, September 11-15, 2000. Revised Lectures (Vol. 1980). Springer Science & Business Media. https://link.springer.com/chapter/10.1007/3-540-45368-7_11#citeas

[3]- Jetty, S., Anbu K, J. P., Jain, P. K., & Hopkinson, A. (2011). OPAC 2.0: towards the next generation of online library catalogues.https://www.researchgate.net/publication/277730139_OPAC_20_towards_the_next_generation_of_online_library_catalogues

[4]- Mandl, T. (2008). Artificial Intelligence for Information Retrieval. In Encyclopedia of Artificial Intelligence (pp. 1-5). IGI Global. doi:10.4018/9781599048499.ch023 The Internet as an agent for social and organizational change | Emerald Insight

4. What are the main research keywords in the field of information retrieval through OPACs for the years 2020-2022?

The methodology:

The methodology of this research aims to review the topic of information retrieval through OPACs using bibliometric analysis. A search was carried out in the SCOPUS database in May 2023 using the following keywords, extracted from the literature and standardised using the SCOPUS Thesaurus:

(("information search" OR "data retrieval" OR "information retrieval" OR "seeking information") AND ("library catalogue" OR "OPAC catalogue" OR "online catalogue")).

This first search produced 328 matching documents. In a second stage, a filtering system was used to refine and reduce the results. This involved selecting the years between 2000 and 2022 and excluding book chapters and conference papers that did not correspond to the research objectives. The focus was on academic articles written in English. The results of the study cover the fields of social sciences, computer science, arts and humanities, excluding other sciences unrelated to the research specialisation. After applying various criteria, the number of documents and research papers was reduced to 95, as shown in Figure 1.

The bibliometric analysis:

The main focus of this research paper is on the topic of information retrieval through OPACs. Given that this topic is not new and is a result of library services' developments and keeping up with technological advancements, libraries have always been at the forefront. This drives us to understand the extent of researchers' interest in this field in light of the advancements in search engines and artificial intelligence techniques over the past two decades. This highlights the need to apply bibliometric analysis, which is a computer-assisted scientific review methodology that can identify key research or authors and their relationships by covering all publications related to a specific topic or field[5]. It is a common and accurate way to explore and analyze large quantities of scientific data. This allows for the detection of fine-grained evolutionary differences in a specific field and sheds light on emerging areas within it[6]. Researchers in the field of bibliometric analysis can track the trajectory of a specific topic by tracing its dissemination across the literature or identify the characteristics of highly downloaded articles to determine their impact[7].

Within this study, data was collected from the SCOPUS database in four sections, including the

[5]- Han, I., Kang, H.-J., Kim, M., & Kwon, G. H. (2020). Mapping the intellectual structure of research on surgery with mixed reality: Bibliometric network analysis (2000–2019). Journal of Biomedical Informatics, 109, 103516. https://doi.org/10.1016/j.jbi.2020.103516 Haut du formulaire

[6]- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

[7]- Ninkov, A., Frank, J. R., & Maggio, L. A. (2022). Bibliometrics: Methods for studying academic publishing. Perspectives on Medical Education, 11(3), 173-176. https://doi.org/10.1007/s40037-021-00695-4

distribution of publications on information retrieval through OPACs in the past two decades, the most relevant and influential journals and authors on this topic, the most active and interested countries in this field, and the employed research keywords and their relationships. The VOSViewer software was employed to build and design bibliometric networks to explore its perspectives for analysis. Data and information were extracted from the documents to conduct this analysis according to the following stages:

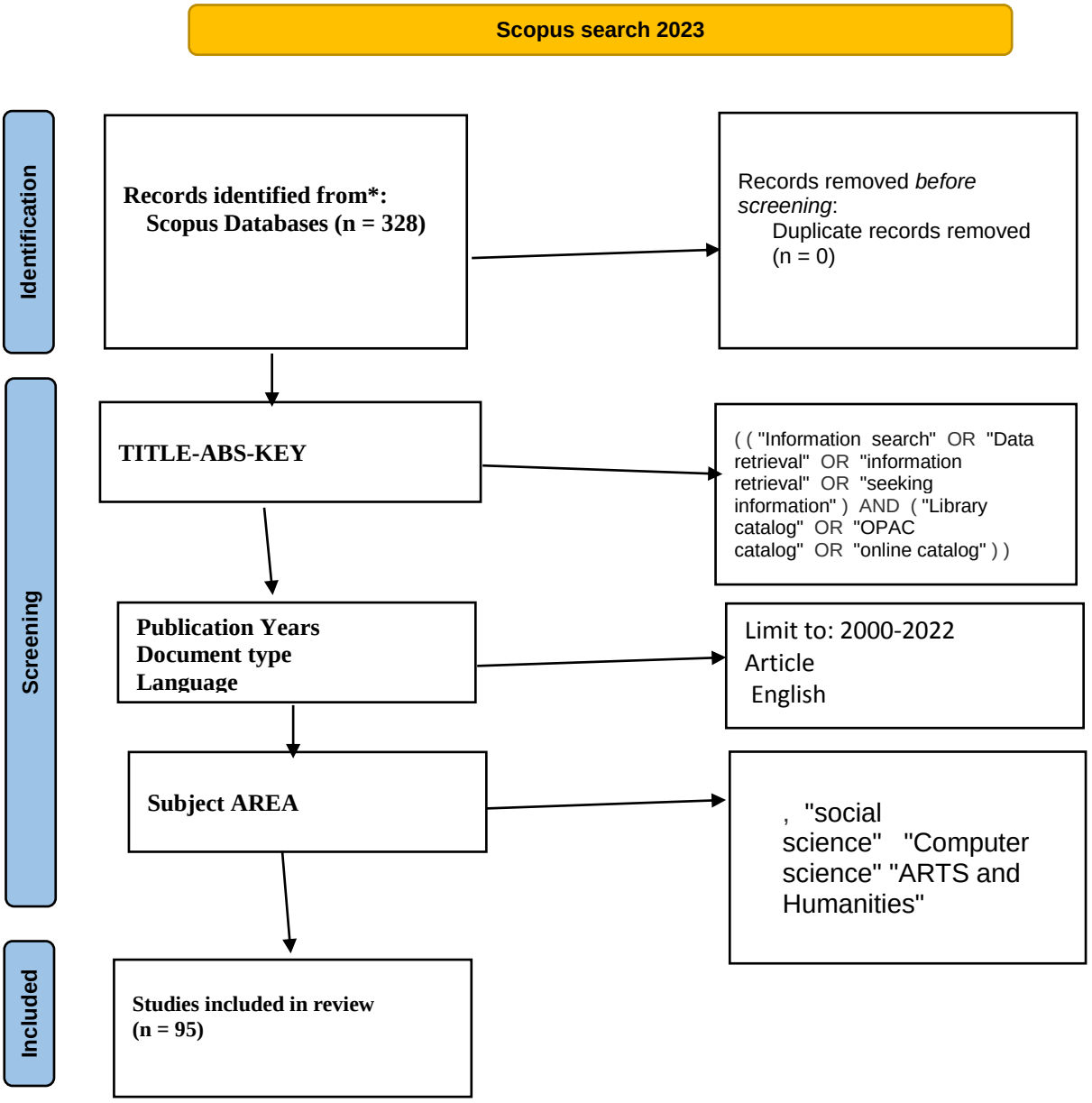


Figure 1: The process involves the search for documents and the filtering of results for bibliometric analysis.

Findings:

The aim of this research paper is to review and quantitatively analyse studies on information retrieval through OPACs using bibliometric analysis over the last two decades. The analysis and review was based on the following research questions:

Research Question 1:

What is the distribution of information retrieval through OPACs from 2020 to 2022?

To answer the first question, an analysis of 95 studies on information retrieval through OPACs from 2000 to 2022 was conducted. The results showed a noticeable fluctuation in the number of published research papers on information retrieval within online library catalogues or OPACs. It was observed that the number of publications has fluctuated over the last two decades. In 2001 a total of 11 papers were published, which then increased and peaked at an average of 10 research papers in 2009. However, from 2010 to 2022 there was a decline and a decrease in the number of publications. It can therefore be concluded that the beginning of the millennium saw a higher number of academic publications compared to the latter part of the millennium, with the number of publications ranging from 5 to 1 research paper. This indicates a shift in research interests and the emergence of new fields over the last two decades. The trend is illustrated in the following figure:

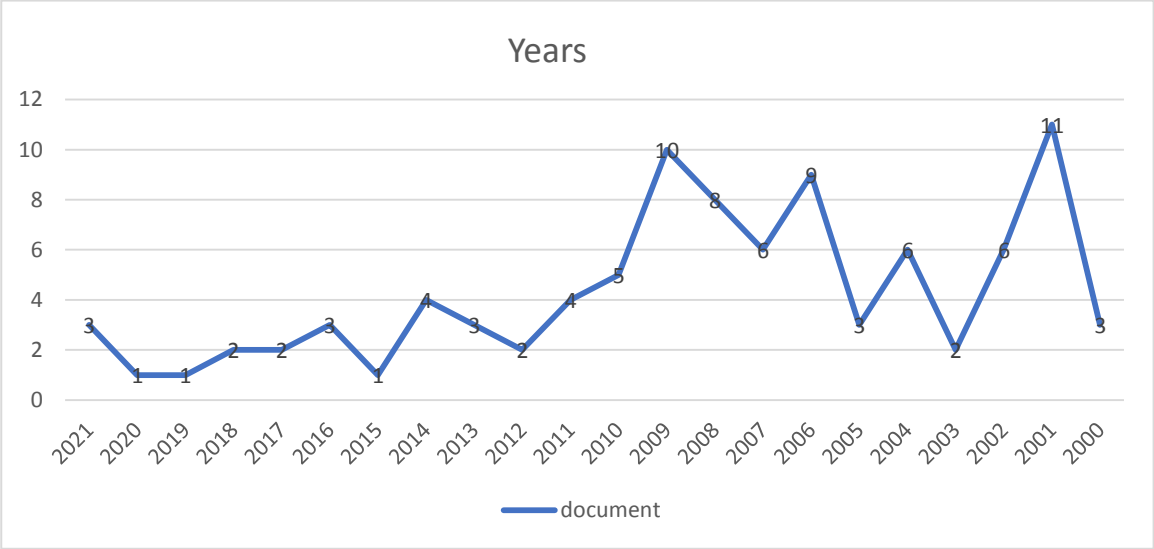


Figure 2: illustrating the distribution of information retrieval through OPACs from 2000 to 2022].

Research question 2:

Which are the most relevant journals and authors for information retrieval via OPACs?

In order to address and analyse the above question, the top 10 domains and 10 authors focused on publishing in the field were selected and processed according to the following criteria "Journal", "Documents", "Total publications", "Total citations", "Cite Score of the journal", "The most cited article", "Cited by" and "Publisher".

Table 1: Top 10 most productive journals in terms of information retrieval via OPACs

Journal	documents	TP	TC	Cite Score	The most cited article	Cited by	Publisher
Cataloging And Classification	14	157	113	0.7	On Overlap and Otherness: A Comparison of Three Vocabularies' Approaches to	4	Cataloging and Classification

Quarterly					LGBTQ+ Identity		Quarterly
Library Hi Tech	13	270	1.237	4.9	Guest editorial: COVID-19 Pandemic and Health Informatics Part 2 *Open Access*	55	Emerald Publishing
Electronic Library	8	210	815	3.9	A quantitative study on utilizing electronic resources to engage children's reading and learning: parents' perspectives through the 5E instructional model *Open Access*	20	Emerald Publishing
Journal Of Documentation	5	334	1.305	3.9	LIS research across 50 years: content analysis of journal articles *Open Access*	25	Emerald Publishing
Journal Of The American Society For Information Science And Technology	4	33	/	/	*Patterns of connections and movements in dual-map overlays: A new method of publication portfolio analysis*	329	Wiley-Blackwell
Online Information Review	4	293	1.585	5.4	Exploring the factors influencing continuous usage intention of academic social network sites	37	Emerald Publishing
Journal Of Web Librarianship	3	43	87	2.0	Implementing a Chatbot on a Library Website *Open Access*	11	Taylor & Francis
OCLC Systems Services International Digital Library Perspectives	3	96	82	0.9	The role of institutional repositories in developing the communication of scholarly research	29	Emerald Publishing

Serials Librarian	3	235	196	0.8	Predatory Publishing: A Catalyst of Misinformation and Disinformation Amongst Academicians and Learners in Developing Countries	06	Taylor & Francis

The previous table (Table 1) illustrates the diversity of journals and the academic impact of each journal, based on factors such as number of publications, citations and other criteria. However, it is noticeable that most of these journals that focus on information retrieval through OPACs are library, information and document-related journals in general. The journal "Cataloging And Classification Quarterly" has the highest productivity in this area with 14 articles, a total of 157 publications and 113 citations. It is followed by "Library Hi Tech" with 13 articles for a total of 270 publications and 1,237 citations. The journal "Electronic Library" has 8 articles with a total of 210 publications and 815 citations. "Journal Of Documentation" has 5 articles with a total of 334 publications and 1,305 citations. "Journal Of The American Society For Information Science And Technology" and "Online Information Review" have 4 articles each, with 33 and 293 publications respectively, and unlisted citations for the former and 1,585 for the latter. "Journal Of Web Librarianship, OCLC Systems Services International Digital Library Perspectives and Serials Librarian have an average of 3 articles each, with total publications of 43, 96 and 235, and total citations of 87, 82 and 169.

Another study was carried out on the researchers most interested in the field of information retrieval through OPACs, taking into account indicators such as country, current affiliation, H-index, total citations (TC), total publications (TP) and the year of publication of the researcher's latest scientific article. This is shown in the table below:

Author	Year of last Publication	TP	TC	H-index	Current Affiliation	Country
Beall, J.	2018	54	1,725	16	University of Colorado Denver, Denver, United States	United States
Fourie, I.	2022	85	532	14	University of Pretoria, Pretoria, South Africa	South Africa
Drabenstott, K.	2013	57	787	17	University of Michigan, Ann	United States

					Arbor, Ann Arbor, United States	
Papadakis, I.	2018	32	99	5	Ionian University, Kerkira, Greece	Greece
Poo, D.C.C.	2022	83	505	10	Info National University of Singapore, Singapore City, Singapore	Singapore
Slone, D.J.	2009	7	171	5	University of South Florida, Tampa, Tampa, United States	United States
Sridhar, M.S.	2004	2	32	2	Indian Space Research Organisation, Bengaluru, India	India
Stefanidakis, M.	2022	33	152	7	Ionian University, Kerkira, Greece	Greece
Tzali, A.	2009	3	10	2	Ionian University, Kerkira, Greece	Greece
Amita	2008	1	21	1	Aligarh Muslim University, Aligarh, India	India

Table 2: Top 10 most productive researchers in the field of OPAC information retrieval

The previous table (Table 1) shows the top 10 most productive authors in the field of OPAC information retrieval. The results of the analysis show the diversity of the researchers' affiliations to different countries and institutions. Researchers from the United States and South Africa have significant and diverse citation impacts, as indicated by their high H-indices, reflecting their strong and influential research contributions. Researchers from Greece and India show growing and

specialised activity in their respective fields.

The researcher (Beall, J.) from the United States, University of Colorado Denver, has the highest productivity in the field, with an H-index of 16 and a total of 1,725 citations. This indicates strong research activity and significant impact in the field until 2018. This is followed by (Fourie, I.) from South Africa, University of Pretoria, with an H-index of 14 and 532 citations, indicating influential contributions and ongoing research activity until 2022. The researcher (Drabenstott, K.) from the United States, University of Michigan, with an H-index of 17 and 787 citations, reflecting a high level of scientific impact until 2013.

The researcher (Papadakis, I.) from Greece, Ionian University, has an H-index of 5 and 99 citations, indicating early and growing research activity until 2018. The researcher (Poo, D.C.C.) from Singapore, National University of Singapore, has an H-index of 10 and 505 citations, indicating significant research activity in this field until 2022. The researcher (Slone, D.J.) from the United States, University of South Florida, has an H-index of 5 and 171 citations, indicating an active research career until 2009.

The researcher (Sridhar, M.S.) from India, Indian Space Research Organisation, has an H-index of 2 and 32 citations, indicating limited but specialised contributions to the field until 2004. The researcher (Stefanidakis, M.) from Ionian University has an H-index of 7 and 152 citations, indicating an increasing impact in the field until 2022.The researcher (Tzali, A.), also from the Ionian University, has an H-index of 2 and 10 citations, indicating limited research activity until 2009.

The researcher (Amita) from India, Aligarh Muslim University, has an H-index of 1 and 21 citations, reflecting an emerging research career until 2008.

Research question 3:

3. Which are the most productive countries in the field of research on information retrieval through OPACs?

In order to analyse the content regarding the most productive countries in the field of information retrieval through OPACs research, the following criteria were selected: "country", "total publications" and "top academic institution", as shown in the following table:

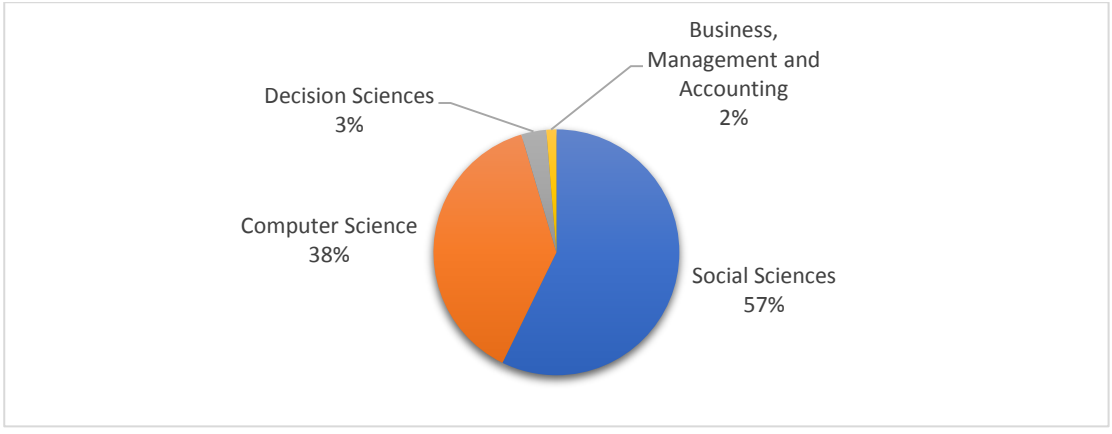
Table 3: Top 10 Countries in Information Retrieval through OPACs Production

country	Education institutions	TP
United States	University of California	55
India	Aligarh Muslim University	7
United Kingdom	University College London	6

Canada	University of Western Ontario	3
Germany	Hamburg University of Applied Sciences	3
Greece		3
Singapore	Nanyang Technological University	3
South Africa	University of Pretoria	3
Spain	University Carlos III de Madrid	3
Australia	Curtin University of Technology	2

In Table 3, for the top 10 most productive countries in the field of information retrieval through OPACs, we observe a significant variation and diversity in the countries and educational institutions interested in publishing in this field. The distribution of productivity is very uneven. The United States leads with a total of 55 publications, which is in line with Table 2 and confirms that most researchers interested in publishing in this field come from the United States. India follows with a total of 7 publications, while the United Kingdom has a total of 6 publications. Greece, Canada, Germany, Sweden and South Africa have a total of 3 publications each. Australia follows with a total of 2 publications. When discussing countries, it is important to consider the specialisations and fields of interest in this research. The majority of research is in the social sciences, with 57%. Computer science follows with 38%, followed by decision sciences with a very low percentage of 3%. Business, management and accounting account for 2%, as shown in the following figure.

Figure 3: Specializations Interested in Publishing within the Studied Field



Research question 4

What are the key research themes for 2020-2022 for information retrieval via OPACs?

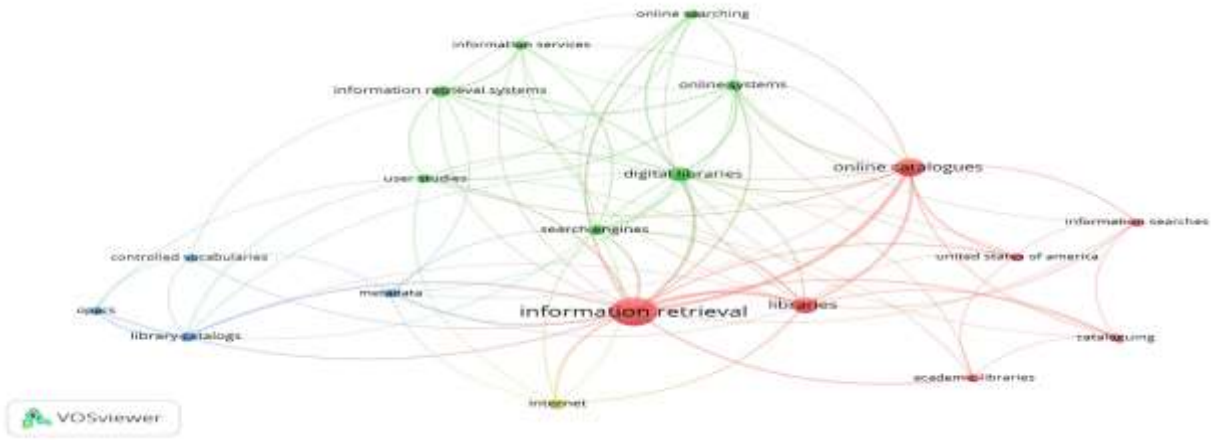


Figure 4: Most frequently used keywords in the analysed scientific publications

For the bibliometric analysis of the most frequently used keywords, 'co-occurrence' was selected as the analysis type using the VOSviewer software, and 'keywords' were marked as the unit. In this context, 19 keywords were identified from the data set, as shown in the following figure.

Through the keyword co-occurrence analysis network, the keywords used in this study were found to be precise and limited compared to the total number of words. They were represented as follows: "information retrieval" = 57 occurrences, "online catalogues" = 23 occurrences, "libraries" = 17 occurrences, "digital libraries" = 15 occurrences, "information retrieval systems" = 9 occurrences, "search engines and online systems" = 8 occurrences, "library catalogues internet" = 7 occurrences. In addition, keywords such as 'user studies', 'online searching', 'metadata', 'information services' were represented as 'occurrences 6'. In addition, keywords such as "United States of America", "Information Searching OPACs", "Controlled Vocabularies", "Cataloguing", "Academic Libraries" were presented as "Occurrences 5".

Furthermore, the analysis and examination of the authors' keywords showed that, among the 265 keywords, the words used in their publications were limited to 14 main keywords. The most frequent keyword was "information retrieval" with "occurrences 45". The most preferred keywords were "online catalogues" with "occurrences 23" compared to "libraries", "cataloguing", "library catalogues", "OPAC", "user studies", "information searching", "Internet", "academic libraries", "digital libraries", "controlled vocabularies" and "metadata", as shown in the following figure.

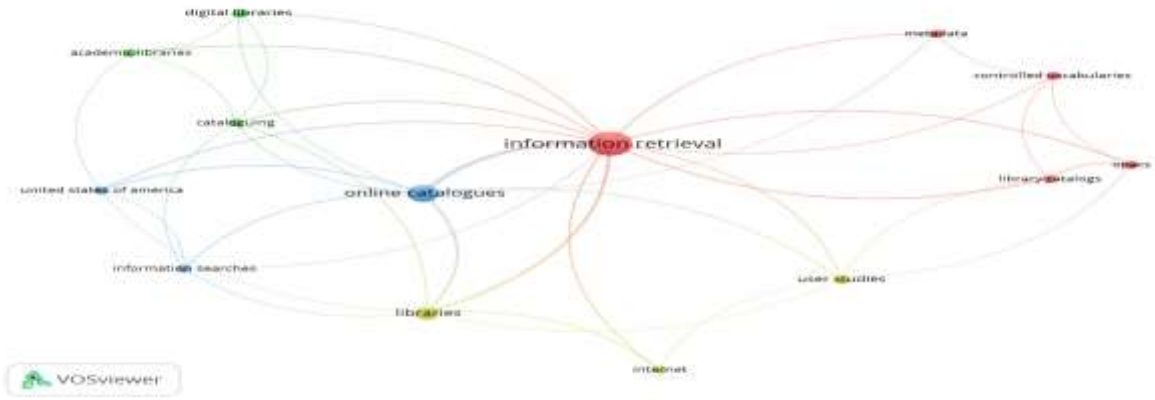


Figure 5: Most frequently used keywords by authors in the field of the study

Discussion:

Following the analysis of a total of 95 research publications collected from the Scopus database, this research review provides an overview of information retrieval through OPACs using quantitative content analysis and bibliometric measurements. The trend analysis shows a fluctuating and low interest in this field over the last decade, as new interests and more complex areas have emerged in the digital environment, such as intelligent scientific search engines and artificial intelligence tools. This analysis also shows that publications on information retrieval through OPACs are mainly focused on narrow fields such as humanities and computing, and are found in journals specialising in information, documents and libraries in general (see Table 1, Figure 2).

The United States of America contributed more than 57% of the literature analysed, with the University of California being the most productive educational institution. The results of the study also show that the majority of researchers and individuals interested in this area of research come from the United States. This is followed by South Africa, Greece, Singapore and India. The primary keywords used were specific and limited, with "information retrieval" being the strongest term. Other related terms include "information searching", "controlled vocabularies" and "metadata" as dimensions of information retrieval. Online catalogues were the most commonly used index term among researchers, as shown in Figure 5.

Conclusion:

The tools and methods of research and information retrieval are constantly evolving as technology advances. From search engines to portals, open repositories, databases and artificial intelligence techniques, these developments have implications for libraries and their online catalogues. Interest in this area has fluctuated in recent years as new and more complex areas of the digital environment have emerged. However, libraries remain an essential pillar for the processing and provision of reliable information and sources, countering the proliferation of predatory publishing found in other tools. This study analysed 95 research papers on information retrieval through OPACs using quantitative methods and bibliometric measures based on the SCOPUS database. The analysis

reflects the varied and limited focus on this area, as well as the specific keywords used. The United States, its journals and researchers showed a greater interest in this field than other countries, including South Africa, Greece, Singapore and India. It should be noted that the results of the study do not necessarily imply a lack of interest on the part of researchers in other regions or disciplines. The scientific journals analysed were predominantly from the library and information field. If other scientific search engines had been included, the results might have been different. Nevertheless, this study could serve as an important reference for researchers to revive this important area within scientific publications.

References:

1. Agosti, M., Crestani, F., & Pasi, G. (Eds.). (2001). *Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000 Varenna, Italy, September 11-15, 2000. Revised Lectures* (Vol. 1980). Springer Science & Business Media. https://link.springer.com/chapter/10.1007/3-540-45368-7_11#citeas
3. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133,* 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070
4. Han, I., Kang, H.-J., Kim, M., & Kwon, G. H. (2020). Mapping the intellectual structure of research on surgery with mixed reality: Bibliometric network analysis (2000–2019). *Journal of Biomedical Informatics, 109,* 103516. https://doi.org/10.1016/j.jbi.2020.103516
5. Jetty, S., Anbu K, J. P., Jain, P. K., & Hopkinson, A. (2011). OPAC 2.0: towards the next generation of online library catalogues. https://www.researchgate.net/publication/277730139_OPAC_20_towards_the_next_generation_of_online_library_catalogues
6. Mandl, T. (2008). Artificial Intelligence for Information Retrieval. In *Encyclopedia of Artificial Intelligence* (pp. 1-5). IGI Global. doi:10.4018/9781599048499.ch023
7. Ninkov, A., Frank, J. R., & Maggio, L. A. (2022). Bibliometrics: Methods for studying academic publishing. *Perspectives on Medical Education, 11*(3), 173-176. https://doi.org/10.1007/s40037-021-00695-4
8. Schwartz, D. (1999), "The Internet as an agent for social and organizational change", *Internet Research*, Vol. 9 No.. https://doi.org/10.1108/intr.1999.17209caa.001