University of Mostaganem-Algeria
VOL: 12 / **N°:** 04 / **(2025)** , p. p. 96/108

International Journal of Social Communication
ISSN: 2437 – 1181
EISSN: 2710 – 8139

Scientometric analysis of cloud-based institutional repositories: Reviewed and Analyzed by VOSViewer from Google Scholar.

Direm Ouisam[1*], Aknouche Nabil[2]

[1] University of Constantine 2 Abdelhamid Mehri, Ouisam.direm@univ-constantine2.dz
Laboratory New Technologies and Their Role in National Development

[2] University of Constantine 2 Abdelhamid Mehri, nabil.aknouche@univ-constantine2.dz
Laboratory New Technologies and Their Role in National Development

,

Received: 07/07/2025 Accepted: 16/12/2025 Published: 31/12/2025

doi 10.53284/2120-012-004-007

Abstract:
As the preservation methods of digital scholarly communication outputs expand , cloud-based institutional repositories have emerged as vital platforms for ensuring long-term accessibility and visibility. The analyzed data, which is drawn from Google Scholar through Harzing's Publish or Perish tool, contain 1000 documents published between 2014 and 2024. The bibliographic analysis conducted the citation network, research domains of the authors, and content of the publications related to the targeted topic.

Keywords: Scientometric analysis , Cloud computing , Institutional repositories, Google scholar database, VOSviewer software, publish perish (Pop).or

*** Corresponding author**

1. INTRODUCTION

Lately, cloud computing has become a revolutionary technology, significantly impacting various sectors, including education, healthcare, and business. It has been coined as an associate degree umbrella term to class of subtle on demand computing services. At the start offered by business suppliers ,like Amazon, Google, and Microsoft. It indicates a model on which a computing infrastructure is viewed as a "cloud," from that businesses and peoples access applications from anywhere in the world on whenever required.

The main principle behind this model is giving computing, storage, and software as a service. Cloud computing offers a transformative infrastructure for IRs, addressing many of the traditional challenges associated with managing and disseminating scholarly outputs. By leveraging cloud services, institutions can overcome limitations in local storage capacity, computing power, and technical expertise.

Despite the growing adoption of cloud-based solutions in academic fields , the intersection between cloud computing and institutional repositories remains neglected in scientometric literature. While there has been a general focus on cloud computing's impact on institutions, the specific challenges and advancements within the academic sphere particularly in managing cloud-based institutional digital repositories. Thus, and given the rapid evolution of cloud technologies and their integration into academic infrastructures, this research aims to fill this gap by exploring the trends and dynamics of cloud computing in institutional repositories from 2014-2024. This study seeks to go through a scientometric analysis of the research literature, identifying key themes, authors, countries, and publication outlets contributing to this field. This study aims to analyze the publication trends, citation patterns, and collaborative networks to explore the evolving role of cloud computing in institutional repositories. It highlights growing interest in cloud-based solutions, with increasing focus on issues like scalability, security, and collaboration. These findings offer valuable insights into current developments and suggest key directions for future research.

2. Literature review:

1. **Gillani, S. M. A. H., Senin, A. B. A., Bode, J., Muniba, & Gillani, S. M. A. H. (2022).** This paper presents a bibliometric analysis of digital entrepreneurial education and student intention, utilizing data from Google Scholar between 2007 and 2022 through Harzing's Publish or Perish tool and VOSViewer. The analysis highlights the relationship among research articles, keywords, and abstracts, focusing on the significance of promoting entrepreneurial intention among university students to address youth unemployment. By mapping citation networks and clustering authors' research areas, the study identifies key trends in entrepreneurial education, revealing 34 authors and 9 clusters of co-authorship. The findings aim to enhance future research methodologies, particularly in developing countries where access to expensive

databases is limited, and emphasize the potential of social entrepreneurship in alleviating poverty and unemployment. Overall, the study seeks to provide a foundational framework for deeper exploration of entrepreneurial intention as a crucial developmental construct. (Dima, 2022)

2.Dima A., Bugheanu A. M., Boghian R. & Madsen D. O. (2022). "this study conducts a complete assessment of scientific production in e-learning and cloud technologies using bibliometric analysis of 637 papers from 2007-2022. The study identifies patterns, frameworks, and the progression of e-learning systems utilizing cloud computing.as a result the study concludes that China is the most producing country, while the most prolific authors originate from Serbia, Japan and Romania. The findings reveal a notable annual growth rate of 20.3% in publications, demonstrating a sustained and rising academic engagement with the topic. The average document age of 6.25 years further suggests that the field is relatively novel and still developing. Proceedings papers dominate the publication types, accounting for 74.09% of the total, which highlights the dynamic and conference-driven nature of research dissemination in this area. The leading subject categories are "Computer Science," "Engineering," and "Educational Research," indicating a multidisciplinary engagement with the topic. In terms of geographic contribution, China emerges as the most productive country, while the most prolific authors are affiliated with institutions in Serbia, Japan, and Romania, reflecting a diverse and internationally distributed research community". (Creto & Grosseck, G. , 2025)

3.S. Karthika, S. Balachandran & S. Sivankalai (2024). "This study provides a comprehensive scientometric evaluation of cloud-based services research, focusing on publications indexed in the Scopus database. The study analyzed 11,321 articles, revealing a significant increase in publications over time, with notable peaks in 2010 and 2018, each accounting for over 1,000 publications. The United States led in publication volume, followed by China and India. The most prolific authors included Rajkumar Buyya and Schahram Dustdar. Top journals in the field were the ACM International Conference Proceeding Series, IEEE Access, and the Journal of Cloud Computing: Advances, Systems and Applications.

The study also identified key research themes, including cloud computing models, service architectures, and security concerns. The findings underscore the growing importance of cloud-based services in various sectors, highlighting the need for ongoing research to address emerging challenges and opportunities in this rapidly evolving field". (Karthika, Balachandran, S., & Sivankalai, S., 2023)

4.Creto D. M., Grosseck G. (2025). "This study examines Romanian Educational Research and its progression towards international recognition over the last thirty years. It examines numerous articles, highlighting both production trends and obstacles, particularly during the COVID-19 pandemic. Conference papers constitute the

predominant output, whilst articles and reviews are very infrequent. It underscores the significance of foreign relationships in augmenting the influence of Romanian research. The report identified five principal areas of emphasis, including digital change in education and teacher preparation.The study primarily examined proceeding papers, articles, and reviews, with proceeding papers identified as the predominant publication type. Approximately one-third of the analyzed papers were open access, with a substantial increase noted in outputs after 2010. The temporal evolution of Web of Science (WOS) publications in Renewable Energy Research (RER) from 1975 to 2024 delineates four distinct phases: the Early Stage, characterized by minimal output; the Growth Stage, marked by a consistent increase in publications; the Peak Production Period, reflecting intensified scholarly activity; and the Diminished and Recovery Phase, indicating a decline followed by a gradual resurgence in research contributions". (M. & Grosseck G., 2025)

2. METHOD AND MATERIALS:

2.1- Data set:

In order to perform a scientometric analysis, Google Scholar was chosen as our data source. "Google Scholar is an online, freely accessible search engine that lets users look for both physical and digital copies of articles. It searches scholarly works from a variety of sources, including academic publishers and universities looking for: Peer-reviewed articles, Theses, Books, Technical reports, Abstracts and Reprints".[i] The time span chosen was 2014-2024. Using the following keywords "cloud computing AND institutional repositorires", (analysis OR trends). We implemented Boolean search operators and we combined keywords with AND, OR to:

- Include multiple concepts (e.g. cloud computing AND institutional repositories)
- Capture synonyms or related terms (e.g. "analysis" OR "trends")

The scientometric analysis workflow consists of a methodical three-phases approach accompanied by extensive visualization. In **phase1** (Preprocessing via Google Scholar), researchers establish search parameters, implement temporal filters (2014-2024), apply inclusion and exclusion criteria, put the keywords using Boolean search operators, gather bibliographic data, and export the findings. This contributes to **phase 2** (Citation Analysis with Publish or Perish), where in the gathered data is imported to compute citation metrics such as h-index and g-index, examine citation trends over time, identify seminal works and authors, and output the processed data in RIS formats. **phase3** (Network Mapping using VOSviewer) involves bringing in the processed data to perform relationship analyses, including co-citation analysis, co-authorship analysis, and keyword co-occurrence analysis, yielding diverse visual maps. The workflow concludes with the Final Output section, which presents four distinct types of visualizations: timeline graphs, bar charts, and evolution graphs.

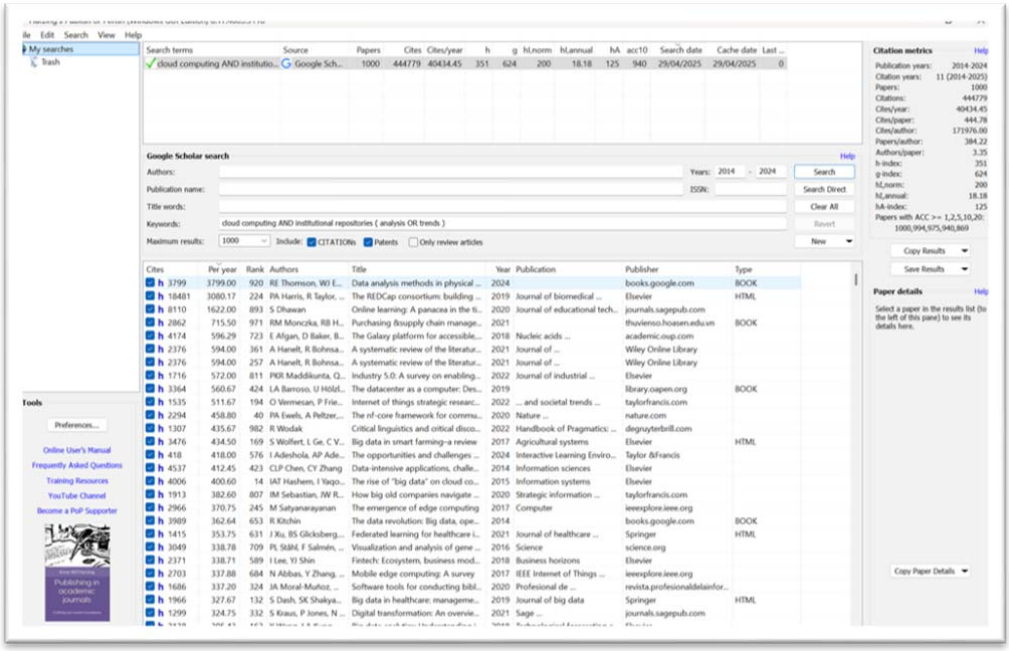
These visual representations provide critical insights into key aspects of the research domain, including emerging trends, the intellectual structure of the field, patterns of collaboration, influential authors, and the thematic evolution over time. Collectively, they contribute to a comprehensive understanding of the scholarly landscape and support informed decision-making for future research directions.

2.2- Pre-processing:

In order to remove unnecessary data and to clean the dataset for more accurate and meaningful analysis, several preprocessing steps are crucial. The report below details the outcomes of the dataset preprocessing. These preprocessing steps included the elimination of duplicate records, correction of inconsistent metadata entries (such as author names and publication titles), normalization of citation formats, and exclusion of incomplete or irrelevant entries. Special attention was given to standardizing journal names and removing entries that lacked essential bibliographic information. Following these steps, the dataset was refined to ensure uniformity and reliability, resulting in a cleaner and more structured collection of records. This enhanced dataset serves as a robust foundation for conducting precise bibliometric analysis, enabling more accurate identification of research trends, citation patterns, and thematic concentrations within the literature

Phase 1: Preprocessing via Google Scholar.

Fig.1. extracting results using advanced research in PoP.



Source: prepared by the authors using publish or purish (PoP) preprocess

As shown in the figure 1 , we inserted the desired period 2014-2024, the keywords were inserted as well. The minimum number of results was 200, we chose the maximum number which is 1000. And the results started to show. In the table below all the results extracted from this advanced research.

Table.1 Citation Metrics Extracted from Publish or Perish Tool

	Value
Publication years	2014 – 2024
Citation years	11 (2014 – 2025)
Papers Metric retrieved	1 000
Total citations	444 779
Citations per year	40 434.45
Citations per paper	444.78
Citations per author	171 976.00
Papers per author	384.22
Authors per paper	3.35
h-index	351
g-index	624
Normalized h-index (h_norm)	200
Annual h-index (h_annual)	18.18
hA-index	125
Papers with ≥ 1 citation	1 000
Papers with ≥ 2 citations	994
Papers with ≥ 5 citations	975
Papers with ≥ 10 citations	940
Papers with ≥ 20 citations	869
ACC(10) (papers with ≥ 10 citations)	940

Note: ACC = Academic Citation Count

Source: prepared by the authors extracted from PoP results.

This search have accrued 444 779 citations—averaging approximately 444.8 citations per paper and 40 434 citations annually.The h-index of 351 and g-index of 624 reveal a substantial nucleus of highly cited publications, while a normalized h-index of 200 and annual h-index of 18.18 reflect sustained citation performance over time. Collaboration is pronounced: with an average of 3.35 authors per paper and 384.22 papers per author, the field exhibits extensive co-authorship networks. Moreover, the fact that 94 % of papers surpass ten citations indicates a

predominantly high-impact literature, collectively signifying the maturation and rapid evolution of research into cloud-based institutional repositories.

Fig 2. chart of citation metrics extracted from PoP.

Source: prepared by the researcher

Table 2. documents types and percentage of documents in open access.

Document Type	Type Count	Type Open Access %
Journal articles	693	14.8%
PDF	48	1.0%
HTML	210	4.5%
BOOK	36	0.8%
Citations	12	0.3%

Source: prepared by authors.

3- RESULTS AND DISCUSSION:

3.1- Documents growth per year:

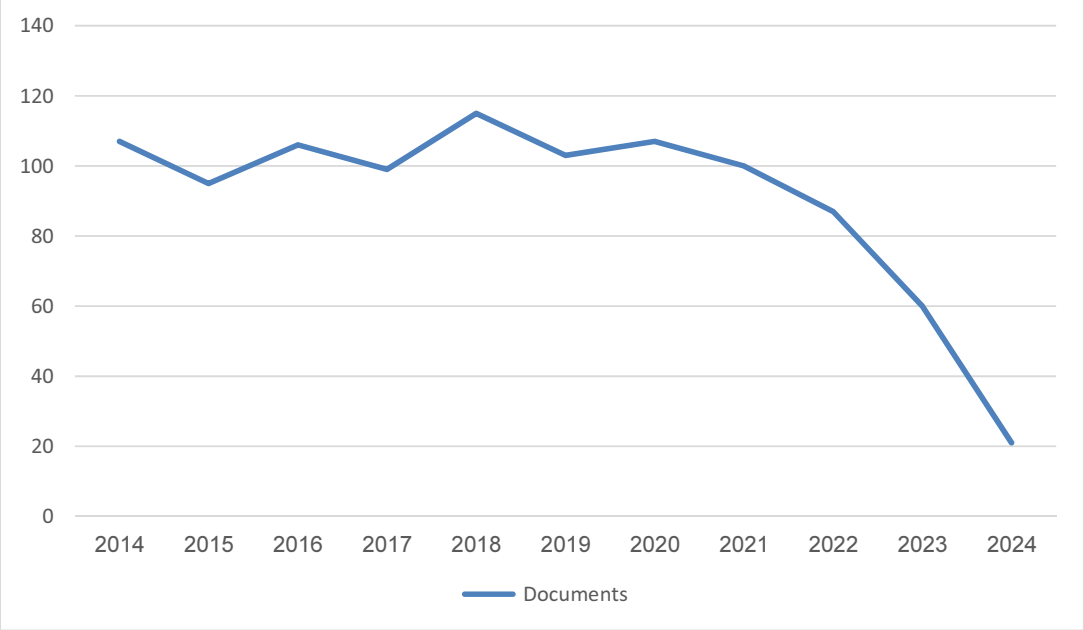
Table 3. documents growth per year using PoP.

Year	Documents
2014	107
2015	95
2016	106
2017	99
2018	115

2019	103
2020	107
2021	100
2022	87
2023	60
2024	21

Source: Prepared by the authors using publish or perish tool.

Fig 3. documents growth per year



Source: prepared by authors using results from PoP in Ris format.

Why Cloud Computing Implementation in Institutional Repositories is Decreasing by Years?

By analyzing the above chart, we can say that cloud computing technology decreased by years. So beetween 2014 and 2020, cloud computing was a hot trend: many institutions were still

exploring and adopting cloud-based repositories. And then by 2021, many repositories had already migrated to the cloud. That means research interests shifted from adoption to optimization or management. So fewer new papers explicitly focused on "cloud computing implementation." As cloud repository systems became more standardized and mature, the need for basic research on implementation naturally declined. However, and by the emergence of New Technologies and from 2020 onward, new technologies like: Artificial Intelligence,Blockchain and Edge Computing competed for attention in academic publications, reducing the focus on traditional cloud computing topics. The Pandemic Disruption (COVID-19) During 2020–2021, research priorities shifted globally due to the pandemic.
A lot of research funding and publication was redirected to healthcare, remote learning, and emergency digital services, slowing down other domains like digital repositories. In addition to Policy and Privacy, Concerns rising data sovereignty laws (like GDPR in Europe) made institutions more cautious about cloud use. Some repositories returned to hybrid models (part-cloud, part-local), leading to lower enthusiasm in publications. (van Eck, 2020)

1. **Top productive authors**:

Table 4. prepared by authors using PoP.

Author	Documents	Total link strength
wang, y	10	3
zhang, y	10	3
li, x	9	2
chang, v	7	1
liu, y	7	0
zeadally, s	7	0
buyya, r	6	0
garcia-peñalvo, fj	6	0
li, j	5	2
wang, l	5	2
wang, j	5	1
anshari, m	5	0
chen, x	5	0
gunasekaran, a	5	0
hanelt, a	5	0
li, z	5	0

Source: prepared by authors from VOSviewer analysis.

1. **Network visualization**

In the network visualization, each item is shown with a label and a circle, whose size reflects the item's weight—the larger the weight, the bigger the label and circle. Some labels may be

hidden to prevent overlap. Colors indicate clusters, and lines represent the strongest links between items (up to 1000 by default). The distance between items suggests their relatedness; closer items are more strongly connected, especially in terms of co-citation.

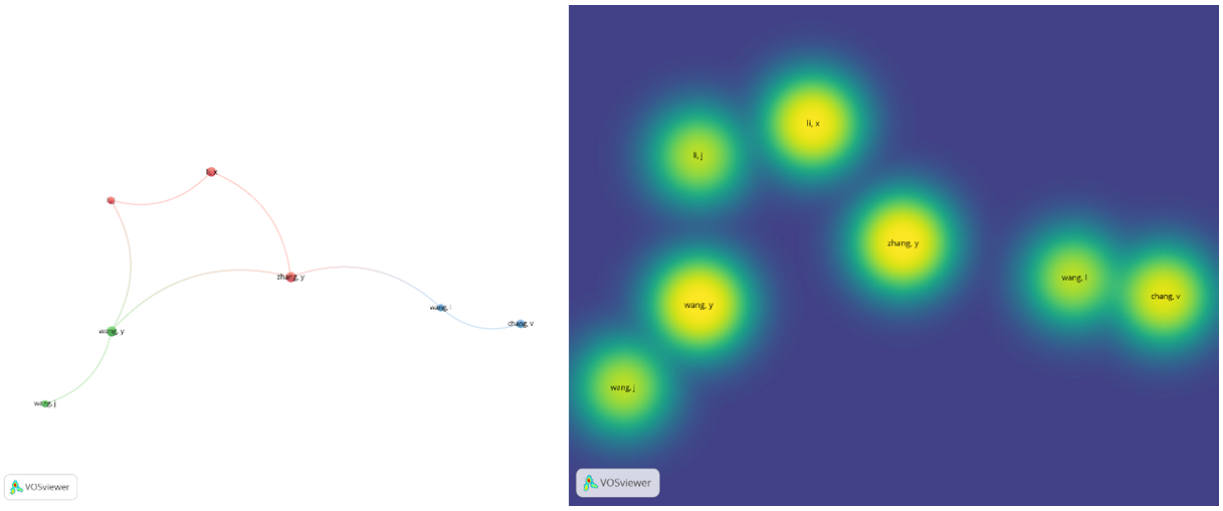
2. **Overlay visualization**

The overlay visualization in VOSviewer functions similarly to the network visualization, although the coloration of things differs. Items may be colored according to their scores, utilizing a gradient from blue (indicating low scores) to yellow (indicating high scores), or they may be assigned user-defined custom colors specified by RGB values in the map file. The overlay visualization will be unavailable in the absence of scores or custom colors.

3. **Density visualization**

VOSviewer offers two categories of density visualizations: item density and cluster density, which may be switched using the respective radio buttons in the settings panel. The item density visualization presents things according to their labels, akin to their representation in the network and overlay visualizations. Every point on the map is color-coded to indicate the density of things in that region. The standard color gradient transitions from blue (representing low density) to green and culminates in yellow (showing high density). A point appears more yellow when surrounded by a bigger quantity of adjacent items with higher weights. In contrast, regions with fewer and less prominent items are shaded nearer to blue. The technical specifics of this visualization's implementation are elaborated over in Van Eck and Waltman (2010). (Library, 2025)

Fig.4. Authors network and density visualization using VOSviewer.



Source: prepared by authors network and density visualization using VOSviewer.

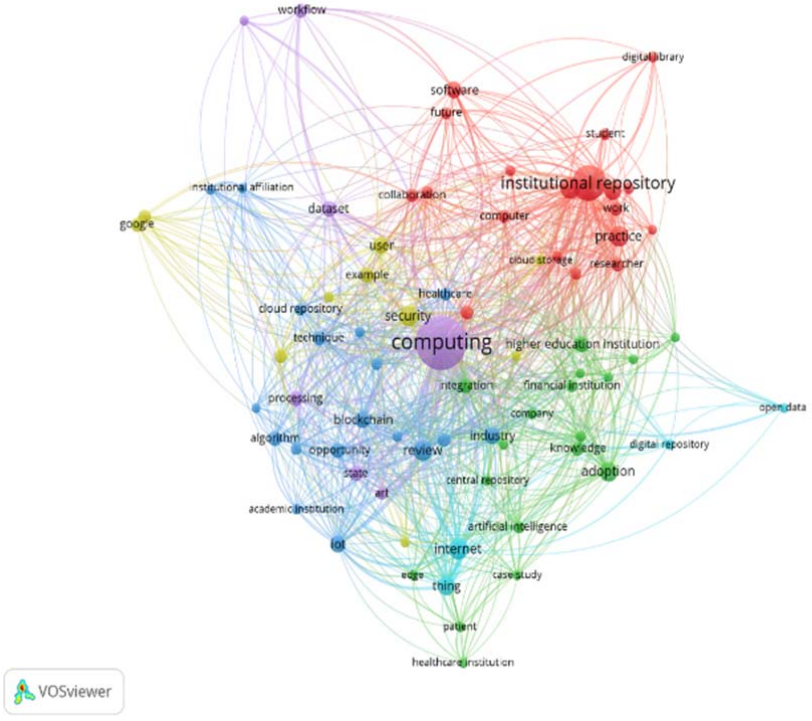
Through table 4 and figure 4 that shows The most productive authors for cloud-based institutional repositories literature. it appears to be **"wang, y"** and **"zhang, y"** who each have 10 documents and a link strength of 3, making them the most productive both in terms of publication count and connections compared to other authors.

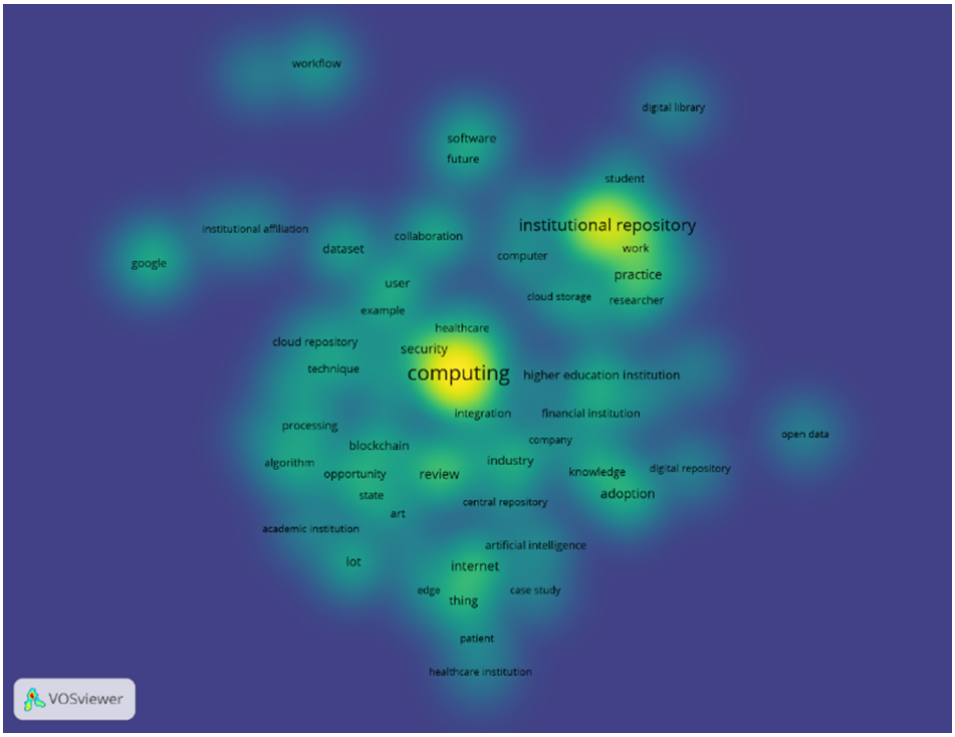
3.3- Top 10 productive publishers:

Rank	Publisher	Number of Documents
1	Elsevier	141
2	Springer	128
3	IEEE Xplore	99
4	MDPI	53
5	Emerald	44
6	Wiley Online Library	44
7	Nature	40
8	Taylor & Francis	34
9	ACM Digital Library	31
10	Google Patents	27

Top keywords related to cloud-based institutional repositories using VOSviewer, VOSviewer tool extracted 5393 terms in title and abstract fields. For each of the 145 terms, a relevance score will be calculated. Based on this score, the most relevant terms will be selected. The default choice is to select the 60% most relevant terms, the number of keywords selected was 87 keywords for 145 term.

Fig. 5 co-citation network and density visualization using VOSviewer





Source: prepared by authors network and density visualization of keywords using VOSviewer.

ii

4. CONCLUSION:

As scientific research on cloud-based institutional repositories continues to grow fueled by progress in technology, academic initiatives, and the increasing demand for digital content management, it becomes essential to understand the current direction of this research field. Identifying emerging trends and the scholarly focus of global research communities provides valuable insight. Therefore, our present scient metric analysis was conducted using the software tool VOS viewer, based on a dataset retrieved from Google Scholar, covering the time span from 2014 to 2024.

the most productive author was wang, y . with a total of 10 top publications to his name, also, the top publisher Elsevier of Research in with 141 total publication.

The findings of this bibliographic analysis underscore a significant and sustained scholarly interest in the role of cloud-based institutional repositories over the past decade (2014-2024). The citation network reveals an increasingly interconnected body of literature, while the diversity of research domains represented indicates a multidisciplinary engagement with the

topic. Moreover, the content of the analyzed publications reflects a growing recognition of the strategic value of cloud technologies in supporting the long-term preservation, accessibility, and global visibility of digital scholarly outputs. These trends suggest that cloud-based repositories are not only evolving as essential infrastructures in the digital knowledge ecosystem but are also shaping the future landscape of research dissemination and academic collaboration.

Bibliography List :

Creto, D. M., & Grosseck, G. . (2025). A bibliometric analysis of Romanian educational research in Web of Science: Trends, challenges, and opportunities for global integration. *Education Sciences*, 358.

Dima, A. B. (2022). Mapping knowledge area analysis in e-learning systems based on cloud computing. Electronics. *Electronics*, 62.

Karthika, S., Balachandran, S., & Sivankalai, S. (2023). A systematic literature analysis of scientometric research in the field of cloud-based services. *Springer*.

Library, H. C. (2025, march 30). . Retrieved from libguides: https://howardcc.libguides.com/c.php?g=52041&p=336275

van Eck, N. J. (2020). *Manual VOSviewer (Version 1.6.15).* leiden: Centre for Science and Technology Studies.