Mapping the Scientific Landscape of Metaverse Using VOSviewer and Bibliometrix

Tayeb Brahimi
Research Energy Lab, College of Engineering
Effat University
Jeddah, Saudi Arabia
(0000-0002-1575-4943) - ORCID

Hala Haneya
Computer Science, College of Engineering
Effat University
Jeddah, Saudi Arabia

***Abstract*— The concept of the Metaverse has gained increasing attention as advances in virtual and augmented reality (AR) technologies have enabled the creation of immersive and interactive virtual environments. However, most of these studies remain independent and only a few studies attempted to investigate their relationships. To analyze key trends in Metaverse research and conduct a thorough bibliometric analysis, we used VOSviewer and Bibliometrix, R-tool package on the Scopus database. Our co-occurrence analysis revealed that the hot topics are related to virtual reality, augmented reality, the Internet of Things, and blockchain, and there are potential areas for future research, such as privacy, security, and education in the Metaverse. In addition, our analysis identified the most active countries and institutions in the field, the top subject areas, as well as potential gaps in the literature that could be explored in future research. This study provides valuable insights into Metaverse research and can help guide future research in this field.**

Keywords*— *Metaverse, virtual reality, augmented reality, Bibliometric, Vosviewer

I. Introduction and Background

The concept of the Metaverse, a virtual universe where users can communicate and interact with each other, has gained significant attention and interest in recent years as virtual and augmented reality (AR) technologies continue to advance and create immersive and interactive virtual environments. This advancement in technology has resulted in an explosion of research in this field, especially since 2020, with numerous studies being published on the potential uses, applications, and challenges of the Metaverse [1], [2]. However, much of the primary Metaverse research is conducted independently and only a few studies have attempted to investigate the relationships and interconnections between different Metaverse research areas [3]. To better understand the scientific landscape of Metaverse research, it is important to conduct a comprehensive data mining, analysis, and visualization of the existing literature. Conducting a comprehensive bibliometric analysis will provide a more complete view of the field, enabling researchers to identify key trends and challenges, as well as active actors and institutions in the area of Metaverse research. According to our literature review only a few documents related to bibliometric analysis and Metaverse. Tilly et al. [4] used bibliometric analysis to study the field of Metaverse in education, revealing research trends, areas of focus, and limitations. The study found a gap in research on lifelogging applications in educational Metaverse and show that the design of Metaverse in education has evolved with a focus on artificial intelligence technologies for a generation. The work by Firmansyah et al. [5] used bibliometric analysis to examine the management issues related to journal publishing in the Southeast Asian context, with a particular emphasis on the Jurnal Pengurusan. Their findings indicate that Malaysia is the most notable country in terms of author affiliation and keyword appearance, and that collaboration among the authors is primarily from the Asian continent. The study provides an evaluation of scholarly literature on management issues in the Southeast Asian context. Baghalzadeh et al. [6] used bibliometric analysis to analyze 1879 academic publications related to Building Information Modeling (BIM), Digital Twins (DT), and the Internet of Things (IoT) in the construction industry. The study identified key publications, conferences, influential authors, countries, organizations, and funding agencies, and suggested potential areas for future research such as the integration of BIM and Metaverse technology, BIM and Artificial Intelligence, and the use of BIM and IoT in the Circular Economy. Zhao et al. [7] used bibliometric analysis to examine the utilization of virtual reality (VR) technology in nursing research. The study results indicate that the United States is the most dominant country in this research field. The main research topics identified are virtual simulation, virtual learning, clinical skills, and dementia. The burst keywords for 2020-2021 were knowledge and simulation. The study also mentions the emergence of the Metaverse concept and suggests that future research in this area will likely increase. Liu et al. [8] conducted a bibliometric analysis to investigate the use of virtual reality (VR) in therapy, focusing on post-traumatic stress disorder (PTSD), anxiety and fear-related disorders (A&F), diseases of the nervous system (DNS), and pain management. The study found that VR-aided therapy is a popular research topic, with the United States being the most dominant country in this field. The research identified four main areas of study in VR-aided therapy. The study also used VOSviewer software to conduct bibliometric analyses on VR-aided therapy from the perspective of Web of Science core collection. The study suggests that future research should focus on integrating more innovative therapies, emphasizing psychological benefits, using game elements, and introducing design research. The study also suggests that the health Metaverse could be a potential application of VR in therapy. Finally, the paper by Abatte et al. [3] used bibliometric analysis to review studies on the concept of the Metaverse. The authors summarize the most significant articles and journals in the field, as well as the most productive and prominent authors, organizations, and countries. The implications of the Metaverse on industry and society are also discussed.

One potential gap in the previous and existing research on the Metaverse is a lack of focus on its potential applications in industries outside of education and healthcare. Besides, while these areas are certainly important and deserving of study, the Metaverse has the potential to impact a wide range of other

979-8-3503-0030-7/23/$31.00 ©2023 IEEE

fields, from entertainment to finance to transportation. Yet, little research has been done to explore the potential applications and trends of the Metaverse. In this paper, we used VOSviewer [9], Bibliometrix, and Biblioshiny platform [10] to analyze trends and perform a bibliometric analysis on Metaverse research using the Scopus database. Data gathered from the Scopus database were selected based on PRISMA model [11]. The paper is organized as follows: Section 2 discusses the research methodology, Section 3 presents the results and discussion, and finally, Section 4 concludes the study.

II. Research Methodology and Data

To effectively analyze a large amount of data collected from a database and investigate relevant research in terms of publication output, research directions, most active countries, and evolution of the field studied, we use bibliometric analysis. The research methodology used in this study is based on a quantitative and qualitative approach using bibliometric analysis and the Visualization of Similarities (VOS) method to understand the intellectual structure of the metaverse and uncover emerging trends in published scientific articles [12], [13].

An important aspect of a bibliometric analysis is the examination of the relationships between different works on the metaverse and the creation of a network diagram to analyze these relationships through citations, co-authorship, or other mappings. In this study, the Scopus database is used for the bibliometric analysis as it is generally considered to have greater publication coverage and uses a system of subject categories and sub-categories to organize its content [14], [15]. The process involves using the advanced search function in the database with additional keywords or criteria.

The output of the bibliometric analysis in the visualization stage includes the number of published papers, the topics of the studies, the extent to which countries collaborated, the power of these relationships, the most popular search terms, the types of scientific journals published, and the number of citations. Additionally, the results can be used to identify gaps in the literature and suggest potential areas for future research.

Figure 1 illustrates the main steps of the analysis process. The first step is to select a database for the bibliometric analysis. The next step is to identify relevant keywords through a thorough literature review and search in the selected database and construct query strings. Using the identified keywords and other criteria, the next step is to refine the search in the database, extract the search results and import them into a CSV file.

It is also important at this stage to perform primary statistical analysis of the data within the database, analyze the data, interpret the results, and draw preliminary conclusions about the topic. The output of the bibliometric analysis in the visualization stage includes the number of published papers, the topics of the studies, the extent to which countries collaborated, the power of these relationships, the most popular search terms, the types of scientific journals published, and the number of citations. Additionally, the results can be used to identify gaps in the literature and suggest potential areas for future research.

The present study employs the Visualization Of Similarities Method (VOS) as implemented in the VOSviewer software version 1.6.18 [9] to analyze, assess, and discover significant research in terms of publications, research trends, active countries and institutions, and the significance and evolution of the Metaverse concept.

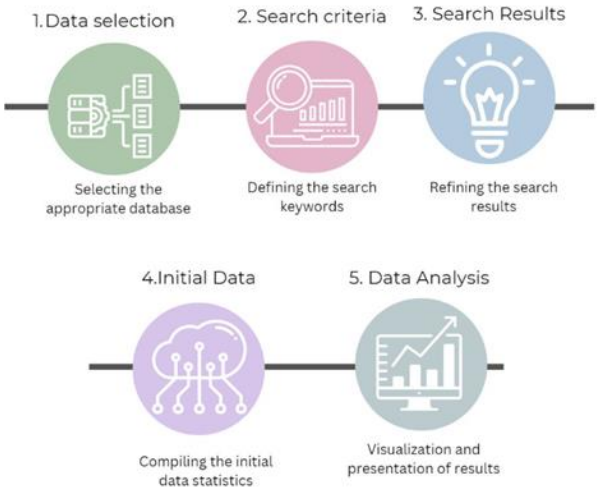


Fig. 1. Main stages of the bibliometric analysis.

To accurately analyze the data collected for this study, it is essential to clean the data and remove duplicates, missing information, and irrelevant documents. We use the PRISMA model (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [16]. By following these guidelines, we can ensure that our data analysis is transparent and reliable. Figure 2 presents the PRISMA flow diagram to illustrate the process of identifying, screening, and determining the eligibility of publications for inclusion in the analysis. The initial search resulted in the extraction of 1054 publications, including articles, conferences, reviews, book chapters, and books. After removing 66 documents that were classified as editorials, letters, or notes, there were 988 remaining documents. These documents were then cleaned by eliminating duplicates and records with missing authors or titles, resulting in a total of 930 eligible documents. Of these, 15 documents were removed because they were not related to the topic under consideration, leaving a total of 915 documents that were included in the bibliometric analysis, as shown in Figure 2.

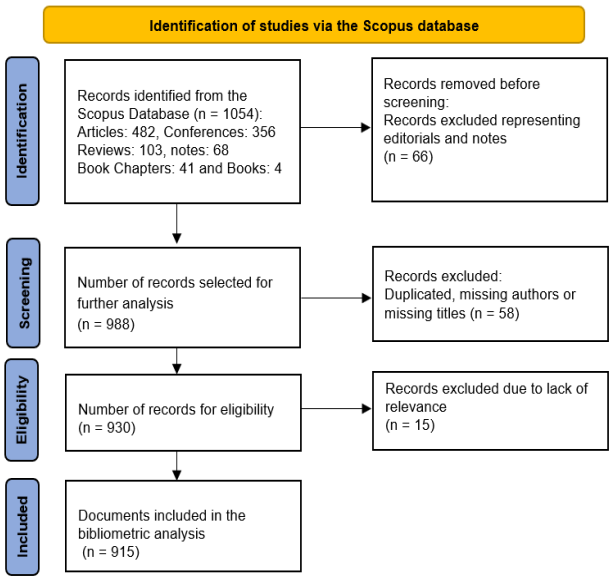


Figure 2. PRISMA flow diagram study selection with identification, screening, eligibility, and inclusion

III. Results and Discussion

The present bibliometric study uses search strings to identify articles related to the Metaverse applied to the title, abstract, and keywords using TITLE-ABS-KEY(("Metaverse*") OR ("Virtual Reality" AND ("Metaverse*")) OR ("augmented reality" AND ("Metaverse*")) OR ("avatars" AND ("Metaverse*")) OR ("IoT" AND ("Metaverse*"))). The analysis was performed on December 10, 2022.

The result stated that 1054 papers had been published; however, after applying the PRISMA model, the number of documents to be evaluated was reduced to 913. Results revealed that from the 913 articles extracted, 80% of them were published from 2020 to 2022 which shows a huge percentage increase. The top five most active countries out of 70 countries who contributed to the publications of documents in Metaverse are China (189), the USA (170), South Korea (109), the United Kingdom (89), and Italy 45 (see Figure 3). The most active institution identified is the Chinese Academy of Sciences, while the top funding agency is the National Natural Science Foundation of China. In terms of total citations, the USA ranks first with 1141 citations, followed by the United Kingdom with 644, South Korea with 528 citations, and China with 335 citations. In terms of subject areas, 31% of the total documents were in computer science, 16% in engineering, 12% in social sciences, 6% in Mathematics, and 5% in business, management, and accounting (see Figure 4).

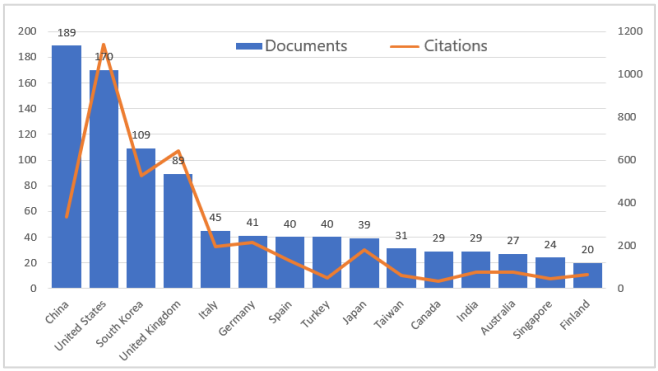


Fig. 3. Number of documents by country (co-authorship).

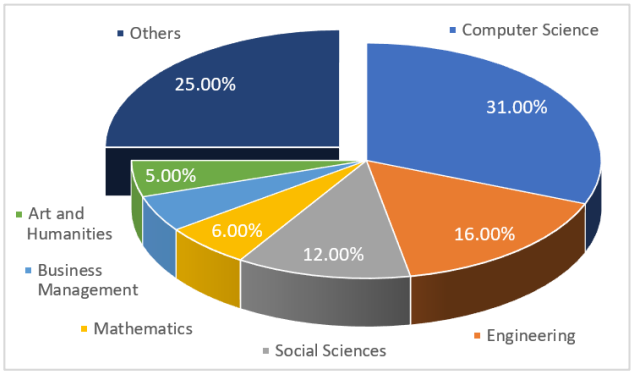


Fig. 4. Distribution of document by subject area.

A. Application of VOSviewer

VOSviewer is a software tool for creating and analyzing networks of bibliographic data, such as that found in Scopus or Web of Science. It is commonly used to create maps of scientific fields, analyze patterns of collaboration, and identify influential authors and papers in a particular domain. VOSviewer is developed by Leiden University Library in the Netherlands and is available as a free download for Windows, Mac, and Linux [9], [17]. In this study, we used VOSviewer version 1.6.18 [9] to analyze bibliometric data and identify patterns and trends. We performed co-occurrence analysis, which involved creating a network with nodes representing words, and edges representing the co-occurrence of the words within the documents. We also conducted co-authorship analysis, which involved creating a network of nodes representing authors and edges representing co-authorship relationships. To compare and combine the data in different ways, we used the overlay visualization feature to display multiple layers of data on the same map. Additionally, we used the density visualization feature to identify areas of the network with a higher concentration of data nodes. Overall, these analyses allowed us to gain insights into the relationships and patterns within the data.

The objective of keyword co-occurrence analysis is to extract frequency data that exist in multiple documents to form a network of relationships among the keywords to identify and highlight the main trend of research in the domain of Metaverse. Figure 5 illustrates the network visualization of author keywords co-occurrence where the size of the circles represents the number of documents published. The larger the circle, the more documents it covers. The distance between two nodes represents the strength of their connection. A lesser distance shows a stronger bond. The thicker the line is, the more often they occur together. The results show that the most frequently used keywords in the Metaverse literature are related to VR, AR, blockchain, and the IoTs (Figure 5).

An analysis of the overlay visualization in figure 6 reveals that recent publications focus on VR, AR, blockchain, and the IoTs, in addition to deep learning, machine learning, and digital twins. The density visualization of the keyword occurrence is illustrated in Figure 7. By overlaying multiple visualizations, it is possible to identify commonalities and differences between the data, making it easier to identify trends and patterns. Each color represents a different density of data, the red color (Metaverse) indicates a high density of data, while the blue color indicates a low density of data.

Both overlay visualization and density visualizations are two useful features used in the VOSviewer software tool for visualizing and analyzing bibliometric data. These findings provided insight into the future Metaverse.

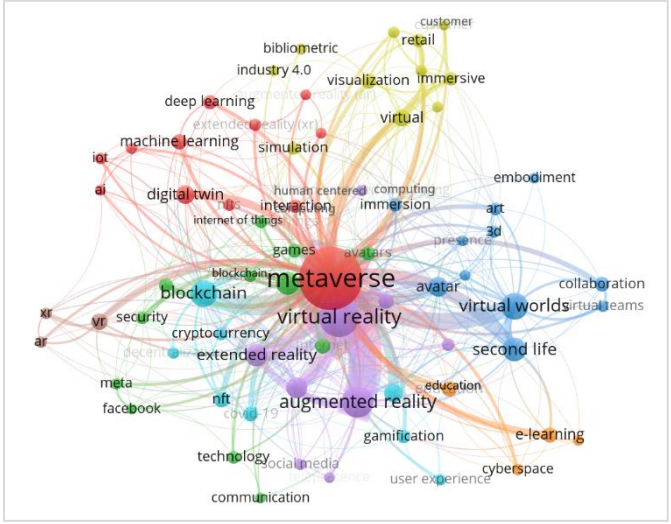


Fig. 5. Network visualization of the author keywords co-occurrence

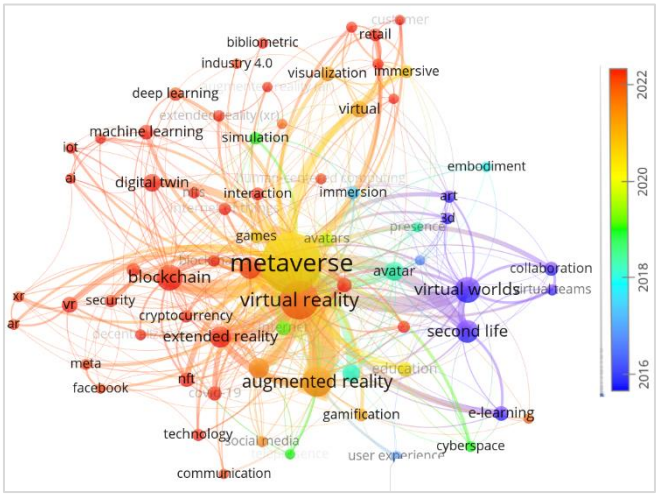


Fig. 6. Overlay Co-occurrence of author keywords.

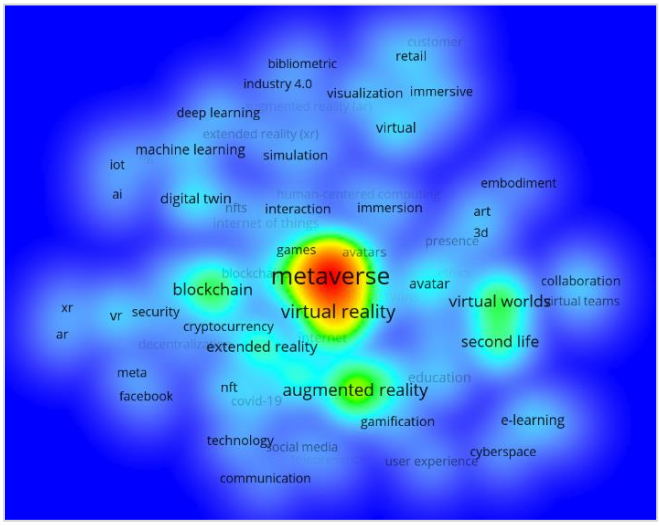


Fig. 7. Density of Co-occurrence of author keywords.

Figure 8 displays the citations by documents. For the 913 documents retrieved and a minimum of 10 citations by document, 81 documents meet the threshold with the three most cited documents by Alanah et al. [18] in 2009 with 215 citations, Dionishio et al. in 2013 [19] with 139 citations and more recently by Sang-Min et al. [20] in 2022 with 98 citations. Alanah et al. paper present a conceptual model for researching Metaverse and their use for virtual team collaboration, This model was designed to facilitate further research and application in virtual teams within the context of Metaverse. The work of Dionishio et al. [19] detailed the necessary advancements that must be made to achieve Metaverse functionality and factors supporting its viability. Finally, the work of Sang-Min et al. present highlights the basic methodologies for used on Ready Player One, Roblox, and Facebook research in games and films. They also outlined limitations for Metaverse as societal influences. Figure 8 suggests that much work is to be done in Metaverse. Furthermore, the collaboration between researchers is weak as most of the nodes are not linked, and only a few researchers are collaborating (colored ones) as shown in the figure.

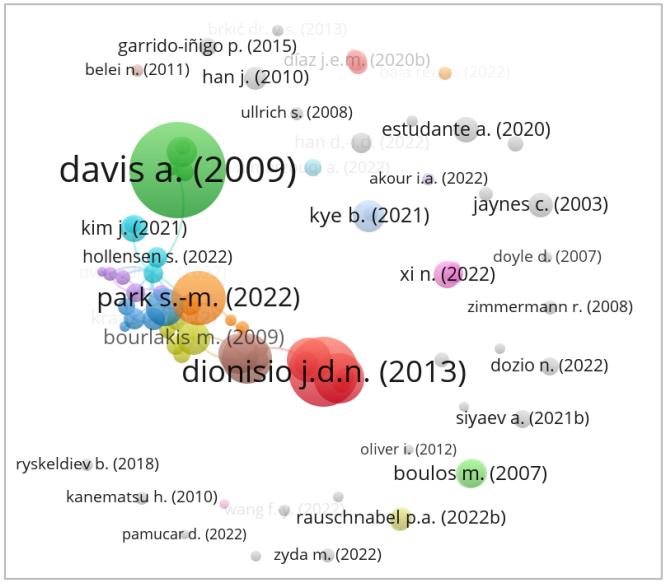


Fig. 8. Citations by documents (min 10 docs).

B. *Application of Bibliometrix*

Bibliometrix is an R-package that provides tools for quantitative research to create comprehensive science mapping and bibliometrics. It is written in R-language, which is an open-source environment and programming language for statistical computing, developed by Massimo Aria and Corrado Cuccurullo [10]. Bibliometrix and its Biblioshiny platform for non-coders provide a complete user-friendly interface to perform extensive bibliometric analysis starting from data collection, followed by data analysis, and ending with data visualization.

A recent survey paper [21] compared various bibliometric analysis tools and found that Bibliometrix "is distinguished by its incorporation of a wide range of different analyses. For instance, it provides analytics and plots at different levels such as author, documents, and sources, in addition to conceptual, intellectual and social structure. In this study, we used Bibliometrix version 4.0.1 which can be run through RStudio, which is an integrated development environment (IDE) for R programming language [22], [23]. RStudio can be run across all operating systems and Bibliometrix can be installed and imported as a package. We started by importing the Bibliometrix package:

```
library(bibliometrix)
```

Next, retrieve the data from the Scopus file, where file is a variable that includes the path to the file:

```
M <- convert2df(file = file, dbsource = "scopus",format = "csv")
```

Then, we applied the three-field function that produces a three-field plot using a Sankey diagram:

```
threeFieldsPlot(M, fields = c("AU", "DE", "DT"), n = c(15, 10, 6))
```

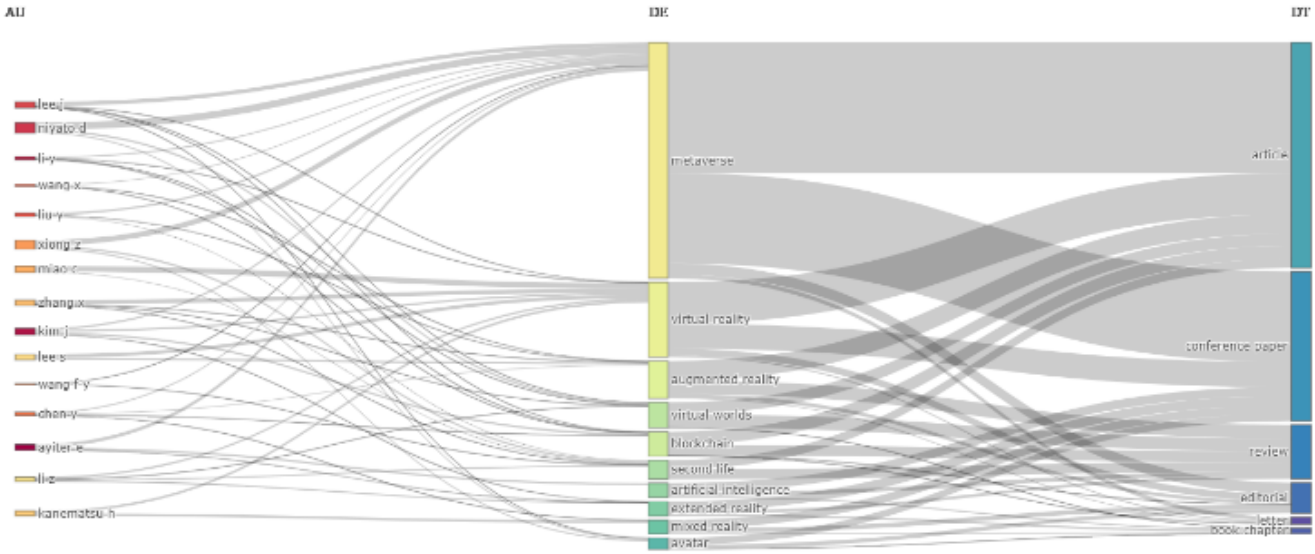
A Sankey diagram illustrates the flow of data from one set to another [24]. It is visualized by nodes, the entities that are being connected, and links, the connections between the nodes. The nodes are represented as rectangles and the links are arrows with widths that are proportional to their importance of it [25]. The three-field function in R takes three parameters, which are the entities to be represented as nodes,

and three integers, that define how many of each entity we will visualize.

Figure 9 shows the resulting three-field plot which visualizes the relationship between the author (AU), keywords (DE), and source type (DT). The number of entities to be visualized has been chosen as 15 authors, 10 keywords, and 6 source types because this combination shows the most relevant information. We find that most of the source types are articles, followed by conference papers. In general, this is expected and matches the numbers found in Figure 2. The dominant keyword is metaverse and this agrees with values in figures 5, 6, and 7.

Metaverse has an incoming flow count of 10 indicating that all authors on the left have used this word in their texts. On the other hand, it has an outgoing flow count of 4 and this means that 4 of the 6 source types have used the word. The two sources that have not been used are review and book chapters. These analyses allowed us to gain a deeper understanding of the interconnections between our main entities – authors, keywords, and sources, within the data.

Overall, the three-field plot in Figure 9 provides valuable insights into the relationships between authors, keywords, and source types in our data. It allows us to see which authors have used which keywords in their texts and which source types have utilized certain keywords. By examining these interconnections, we can gain a more comprehensive understanding of the data and the themes that emerge within it.



,

Fig. 9. Three-Field Plot visualizing the relation between authors (AU), Keywords (DE), and Source Type (DT).

IV. Conclusion

The bibliometric analysis conducted using Bibliometrix and VOSviewer showed that the topic of the metaverse has received substantial attention in recent years, with the majority of articles on the topic being published within the past two years. The top countries in terms of publication were China, the USA, South Korea, the United Kingdom, and Italy. The most active institutions in terms of publications were the Chinese Academy of Sciences and the National Natural Science Foundation of China. In terms of total citations, the USA was the top contributor, followed by the United Kingdom, South Korea, and China. The most common subject areas for articles on the metaverse were computer science, engineering, and social sciences, with keywords related to virtual reality (VR), augmented reality (AR), blockchain, and the Internet of Things (IoT) being the most frequently used. Additionally, recent publications have also focused on machine learning, deep learning, and digital twins.

The field plot visualization was a useful tool for identifying trends and patterns in the data and understanding the connections between different entities within the dataset. However, it is important to note that this analysis may not capture all forms of research outputs on the metaverse, such as workshops and other types of publications. Despite this limitation, the analysis still provides valuable insights into the state of the art in metaverse research and can help guide future research in this field by identifying patterns, trends, and potential gaps in the literature.

Acknowledgment

The authors acknowledge the support provided by the College of Engineering at Effat University, Jeddah, Saudi Arabia.

REFERENCE

[1] M. Vasarainen, S. Paavola, and L. Vetoshkina, "A Systematic Literature Review on Extended Reality : Virtual, Augmented and Mixed Reality in Working Life," Oct. 2021 [Online]. Available: https://helda.helsinki.fi/handle/10138/337180. [Accessed: Dec. 10, 2022]

[2] Y. K. Dwivedi *et al.*, "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *International Journal of Information Management*, vol. 66, p. 102542, Oct. 2022, doi: 10.1016/j.ijinfomgt.2022.102542.

[3] S. Abbate, P. Centobelli, R. Cerchione, E. Oropallo, and E. Riccio, "A first bibliometric literature review on Metaverse," in *2022 IEEE Technology and Engineering Management Conference (TEMSCON EUROPE)*, Apr. 2022, pp. 254–260, doi: 10.1109/TEMSCONEUROPE54743.2022.9802015.

[4] A. Tlili *et al.*, "Is Metaverse in education a blessing or a curse: a combined content and bibliometric analysis," *Smart Learning Environments*, vol. 9, no. 1, p. 24, Jul. 2022, doi: 10.1186/s40561-022-00205-x.

[5] E. A. Firmansyah, H. Wahid, A. Gunardi, and F. A. Hudaefi, "A Scientometric Study on Management Literature in Southeast Asia," *Journal of Risk and Financial Management*, vol. 15, no. 11, p. 507, Nov. 2022, doi: 10.3390/jrfm15110507.

[6] M. Baghalzadeh Shishehgarkhaneh, A. Keivani, R. C. Moehler, N. Jelodari, and S. Roshdi Laleh, "Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in Construction Industry: A Review, Bibliometric, and Network Analysis," *Buildings*, vol. 12, no. 10, p. 1503, Oct. 2022, doi: 10.3390/buildings12101503.

[7] J. Zhao, Y. Lu, F. Zhou, R. Mao, and F. Fei, "Systematic Bibliometric Analysis of Research Hotspots and Trends on the Application of Virtual Reality in Nursing," *Front Public Health*, vol. 10, p. 906715, 2022, doi: 10.3389/fpubh.2022.906715.

[8] Z. Liu, L. Ren, C. Xiao, K. Zhang, and P. Demian, "Virtual Reality Aided Therapy towards Health 4.0: A Two-Decade Bibliometric Analysis," *Int J Environ Res Public Health*, vol. 19, no. 3, p. 1525, Jan. 2022, doi: 10.3390/ijerph19031525.

[9] VOSviewer, "Visualizing scientific landscapes," *VOSviewer*, 2022. [Online]. Available: https://www.vosviewer.com//. [Accessed: Sep. 06, 2022]

[10] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, Nov. 2017, doi: 10.1016/j.joi.2017.08.007.

[11] PRISMA, "PRISMA," 2020. [Online]. Available: https://www.prisma-statement.org/. [Accessed: Dec. 03, 2022]

[12] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, pp. 285–296, Sep. 2021, doi: 10.1016/j.jbusres.2021.04.070.

[13] N. Donthu and S. Kumar, "Forty-five years of Journal of Business Research: A bibliometric analysis," *Journal of Business Research*, vol. 109, pp. 1–14, Mar. 2020, doi: 10.1080/0907676X.2021.2004177.

[14] S. Beatty, "Five steps to creating a citation overview in Scopus," *Scopus*, 2001. [Online]. Available: https://blog.scopus.com/posts/five-steps-to-creating-a-citation-overview-in-scopus. [Accessed: Dec. 03, 2022]

[15] R. Pranckutė, "Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World," *Publications*, vol. 9, no. 1, p. 12, Mar. 2021, doi: 10.3390/publications9010012.

[16] M. J. Page, D. Moher, and J. E. McKenzie, "Introduction to PRISMA 2020 and implications for research synthesis methodologists," *Research Synthesis Methods*, vol. 13, no. 2, pp. 156–163, 2022, doi: 10.1002/jrsm.1535.

[17] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010, doi: 10.1007/s11192-009-0146-3.

[18] D. Alamah, M. John, O. Dawin, K. Deepak, and Z. Ilze, "'Avatars, People, and Virtual Worlds: Foundations for Research in Metav' by Alanah Davis, John Murphy et al.," *ournal of the Association for Information Systems*, vol. 10, no. 2, 2009, doi: DOI: 10.17705/1jais.00183. [Online]. Available: https://aisel.aisnet.org/jais/vol10/iss2/1/. [Accessed: Dec. 17, 2022]

[19] J. D. N. Dionisio, W. G. B. III, and R. Gilbert, "3D Virtual worlds and the metaverse: Current status and future possibilities," *ACM Comput. Surv.*, vol. 45, no. 3, p. 34:1-34:38, Jul. 2013, doi: 10.1145/2480741.2480751.

[20] S.-M. Park and Y.-G. Kim, "A Metaverse: Taxonomy, Components, Applications, and Open Challenges," *IEEE Access*, vol. 10, pp. 4209–4251, 2022, doi: 10.1109/ACCESS.2021.3140175.

[21] J. A. Moral-Muñoz, E. Herrera-Viedma, A. Santisteban-Espejo, and M. J. Cobo, "Software tools for conducting bibliometric analysis in science: An up-to-date review," *Profesional de la información*, vol. 29, no. 1, Jan. 2020, doi: 10.3145/epi.2020.ene.03. [Online]. Available: https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2020.ene.03. [Accessed: Dec. 29, 2022]

[22] "Posit | The Open-Source Data Science Company." [Online]. Available: https://posit.co/. [Accessed: Dec. 29, 2022]

[23] "RStudio IDE (Integrated Development Environment) - R for the Social Sciences." [Online]. Available: https://hbs-rcs.github.io/R_Intro-gapminder/rstudio-ide/. [Accessed: Dec. 29, 2022]

[24] "Sankey Diagram | Charts | Google Developers." [Online]. Available: https://developers.google.com/chart/interactive/docs/gallery/sankey. [Accessed: Dec. 29, 2022]

[25] "Sankey diagram – from Data to Viz." [Online]. Available: https://www.data-to-viz.com/graph/sankey.html. [Accessed: Dec. 29, 2022]