PENINSULA PUBLISHING LLC

SHIFRA
Vol. **(2025)**, 2025, **pp**. 46–61
ISSN: 3078-3186

Research Article

Federated Learning in Healthcare: A Bibliometric Analysis of Privacy, Security, and Adversarial Threats (2021-2024)

Mohammed Amin Almaiah[1,*], Rejwan Bin Sulaiman[2], Umar Islam[3], Youakim Badr[4], Fuad Ali El-Qirem[5]

[1] *King Abdullah the II IT School, The University of Jordan, Amman 11942, Jordan*

[2] *School of Computer science and Technology, Northumbria University, Newcastle Upon Tyne, UK*

[3] *Department of Computer Science, IQRA National University, Swat Campus, Peshawar, Pakistan*

[4] *Computer Science and Engineering Department, The Pennsylvania State University, Malvern, PA, USA*

[5] *Faculty of Architecture and Design, Al-Zaytoonah University of Jordan, Amman 11733, Jordan*

ARTICLE INFO

Article History
Received 12 Oct 2024
Revised: 2 Dec 2024
Accepted 2 Jan 2025
Published 17 Jan 2025

Keywords
Federated Learning,
Bibliometric Analysis,
Healthcare,
Data Privacy,
Adversarial Attacks.

ABSTRACT

Federated Learning (FL) has rapidly emerged as a transformative machine learning approach, enabling healthcare institutions to collaboratively build predictive models without compromising patient data privacy. As healthcare increasingly adopts digital technologies, federated learning offers promising solutions to critical issues such as data privacy, security, data poisoning, and adversarial attacks. Despite the recognized potential of FL, significant gaps persist in existing research, particularly concerning comprehensive security frameworks and practical healthcare applications. This bibliometric analysis systematically explores the research landscape from 2021 to 2024, explicitly focusing on data privacy, security threats, and adversarial attacks within federated learning in healthcare. Utilizing bibliometric data from the Scopus database, the study identifies key thematic trends, evaluates global collaborative networks, and assesses contributions from leading institutions and countries. Findings reveal rapidly growing scholarly interest, robust international collaboration, and notable institutional contributions, with a specific emphasis on privacy-preserving techniques, healthcare-specific applications, and emerging technologies such as blockchain and edge computing. The analysis also highlights critical limitations due to incomplete bibliographic metadata. This research provides a comprehensive understanding of current trends and identifies future directions to enhance the security and privacy framework of federated learning in healthcare.

1. INTRODUCTION

Bibliometric analysis, a robust and systematic method for evaluating scientific literature, offers significant value in understanding the landscape, trends, and research impacts within a specific field. By quantitatively analyzing publication patterns, citation networks, collaborative dynamics, and thematic trends, bibliometric analyses help researchers and policymakers identify gaps, emerging topics, and influential contributors within a research domain [9,10].

In recent years, healthcare has seen an exponential rise in the adoption of digital technologies, leading to an unprecedented accumulation of patient data. While this data explosion offers tremendous opportunities for advancing healthcare through data-driven insights, it simultaneously raises significant concerns regarding data privacy, security, and integrity. Federated learning addresses these challenges by allowing collaborative model training across distributed datasets while ensuring data remains locally stored and secure.

The importance of federated learning in healthcare extends far beyond academic interest. It directly impacts patient care quality, facilitates international collaboration, and addresses pressing ethical concerns surrounding data security. As digital health solutions become globally pervasive, addressing privacy and security threats such as data poisoning and adversarial attacks becomes not only desirable but imperative.

*Corresponding author email: m.almaiah91@gmail.com
DOI: https://doi.org/10.70470/SHIFRA/2025/002

This Federated Learning (FL), a revolutionary approach in machine learning that enables models to learn from decentralized datasets without compromising data privacy [1-3], has rapidly emerged as a critical solution in healthcare technology [4-5]. Imagine a world where hospitals and medical institutions can collaboratively train predictive models without ever sharing sensitive patient data, significantly reducing privacy risks. This is precisely what federated learning promises enhanced collaboration, improved healthcare outcomes, and stringent privacy protection [6-8].

1.1 Problem Statement

Despite the clear advantages and potential of federated learning in healthcare, existing approaches are not without limitations. Current research faces critical gaps, particularly concerning vulnerabilities like data poisoning, adversarial attacks, and persistent privacy threats. Most conventional approaches inadequately address the sophisticated security risks associated with federated learning frameworks, especially in sensitive healthcare environments. Furthermore, while many studies emphasize theoretical aspects of federated learning, practical applications and comprehensive security frameworks remain underexplored [11-14].

Given the high stakes involved in healthcare data management, the limitations of current methodologies present significant risks. The failure to adequately protect data integrity and privacy could undermine trust, compromise patient confidentiality, and ultimately jeopardize healthcare outcomes. Therefore, addressing these gaps is not just necessary but urgent.

1.2 Research Objectives and Questions

This study aims to systematically analyze the evolving research landscape of federated learning in healthcare, focusing explicitly on issues related to data poisoning, privacy, and security. Specifically, the objectives of this research are to:

- Identify major thematic clusters and trends within federated learning research in healthcare.
- Examine the global collaborative networks and research productivity trends.
- Evaluate the scientific contributions and impact of leading institutions and countries in the domain.

To achieve these objectives, this study addresses the following research questions:

1. What are the dominant research themes and focal areas in federated learning related to healthcare security and privacy?
2. Which institutions and countries lead research in federated learning in healthcare, and how do their contributions evolve over time?
3. How do international collaborations influence the development and implementation of federated learning technologies in healthcare?

1.3 Scope and Limitations

This research focuses explicitly on bibliometric data from the Scopus database, covering publications from 2021 to 2024. It encompasses studies addressing federated learning in healthcare, specifically emphasizing issues of data poisoning, adversarial attacks, privacy, and security.

Despite its comprehensive nature, the study acknowledges certain limitations. Primarily, there are notable gaps in bibliographic metadata, particularly the complete absence of cited references and science categories, limiting a deeper citation-based and thematic network analysis. Additionally, this analysis relies solely on publications indexed in Scopus, potentially excluding relevant studies indexed elsewhere.

2. METHODOLOGY

This study employs a bibliometric analysis to systematically explore research trends and patterns in the domain of Federated Learning (FL) in healthcare, specifically emphasizing issues of data poisoning, privacy, and security. The analysis utilizes the Scopus database, chosen for its comprehensive coverage, extensive citation metrics, and global recognition as a reliable source for bibliometric analyses. Scopus provided a broad spectrum of literature, offering diverse insights into the current research landscape.

2.1 Search Strategy

A detailed and structured search query was developed to capture relevant studies published between 2021 and 2024. The query targeted the title, abstract, and keywords fields in the Scopus database, formulated as follows:

("Federated Learning" OR "FL") AND ("Healthcare" OR "Medical") AND ("Data Poisoning" OR "Adversarial Attack" OR "Privacy" OR "Security")

Applying this query yielded a total of 2113 documents. These documents formed the initial corpus for the subsequent analysis.

2.2 Inclusion and Exclusion Criteria

To ensure relevance and consistency, the following inclusion and exclusion criteria were applied:

- Inclusion Criteria:
  - Publications specifically addressing Federated Learning in the context of healthcare or medical applications.
  - Studies published from 2021 to 2024.
  - Articles explicitly discussing or addressing issues of data poisoning, adversarial attacks, privacy, or security.
- Exclusion Criteria:
  - Publications unrelated to healthcare or medical contexts.
  - Studies published before 2021.
  - Papers without full-text availability or lacking abstracts.
- Study Selection
  The retrieved documents were carefully reviewed based on titles, abstracts, and keywords to ensure adherence to the defined inclusion criteria. After initial screening and refinement, the dataset comprising 2113 documents was finalized for analysis. The analysis was conducted using the R programming language and RStudio software, employing the biblioshiny package, renowned for its robust visualization and bibliometric analytics capabilities.
- Completeness of Bibliographic Metadata
  During the extraction and analysis phase, metadata completeness emerged as a notable issue. Certain metadata fields were incomplete or entirely missing, as summarized in Table I.

TABLE I COMPLETENESS ASSESSMENT OF BIBLIOGRAPHIC METADATA

Metadata	Description	Missing Counts	Missing %	Status
AU	Author	0	0.00	Excellent
DT	Document Type	0	0.00	Excellent
SO	Journal	0	0.00	Excellent
LA	Language	0	0.00	Excellent
PY	Publication Year	0	0.00	Excellent
TI	Title	0	0.00	Excellent
TC	Total Citation	0	0.00	Excellent
AB	Abstract	7	0.33	Good
C1	Affiliation	83	3.93	Good
DI	DOI	137	6.48	Good
DE	Keywords	316	14.96	Acceptable
ID	Keywords Plus	388	18.36	Acceptable
RP	Corresponding Author	619	29.29	Poor
CR	Cited References	2113	100.00	Completely missing
NR	Number of Cited References	2113	100.00	Completely missing
WC	Science Categories	2113	100.00	Completely missing

Given these metadata limitations, especially the complete absence of cited references and science categories, the analysis did not depend exclusively on these missing fields. Instead, the study focused on reliably available metadata, including author details, document type, publication year, titles, abstracts, and keywords. This approach ensures accuracy and robustness in presenting the research trends despite the noted limitations in the metadata completeness.

3. INFORMATION ABOUT DATA

The bibliometric dataset analyzed in this study spans from 2021 to 2024 and includes 2113 documents sourced from a diverse set of 907 journals, books, and other publications, reflecting an impressive annual growth rate of 84.93%. The average age of the documents is relatively recent, approximately 1.84 years, indicating the emerging nature of research on Federated Learning (FL) in healthcare. The dataset demonstrates a substantial research impact, averaging 13.22 citations per document, indicative of strong scholarly engagement and the topic's significance.

Within the analyzed documents, there were 7233 occurrences of "Keywords Plus" and 3478 occurrences of author-defined keywords, underscoring the diversity of research topics explored. A total of 6516 authors contributed to the corpus, though only 52 authors published single-authored documents, reflecting the highly collaborative nature of this research area. Indeed, the dataset featured an average of 4.75 co-authors per document, with 36.02% international co-authorship, illustrating significant global collaboration and interdisciplinary research efforts.

Document types within the corpus predominantly included conference papers (917) and articles (836), followed by book chapters (135) and reviews (96). Less common document formats such as editorials, notes, and short surveys were minimally represented, highlighting a clear preference toward peer-reviewed research articles and conference proceedings. This distribution of document types emphasizes the active and ongoing scholarly discourse in federated learning concerning healthcare data security, privacy concerns, and adversarial threats, table II summarize this information.

TABLE II MAIN INFORMATION ABOUT DATA

Description	Results
Timespan	2021:2024
Sources (Journals, Books, etc)	907
Documents	2113
Annual Growth Rate %	84.93
Document Average Age	1.84
Average citations per doc	13.22
References	1
DOCUMENT CONTENTS	
Keywords Plus (ID)	7233
Author's Keywords (DE)	3478
AUTHORS	
Authors	6516
Authors of single-authored docs	52
AUTHORS COLLABORATION	
Single-authored docs	131
Co-Authors per Doc	4.75
International co-authorships %	36.02
DOCUMENT TYPES	
article	836
article article	1
article book	1
article book chapter	1
article conference paper	2
article conference review	1
article review	1

book	16
book article	1
book chapter	135
book chapter article	1
book conference paper	1
conference paper	917
conference paper article	2
conference paper book chapter	1
conference paper conference paper	4
conference paper review	1
conference review	78
conference review conference paper	1
data paper	2
editorial	4
erratum	1
letter	1
note	4
review	96
review article	2
short survey	2

4. ANNUAL SCIENTIFIC PRODUCTION

The analysis of annual scientific production between 2021 and 2024 (see Figure 1) revealed a significant upward trajectory in research output related to Federated Learning in healthcare. Starting with 154 published articles in 2021, the production more than doubled to 338 articles in 2022. The trend continued, nearly doubling again to reach 647 articles in 2023. By 2024, the scientific production peaked dramatically at 974 articles, clearly illustrating the rapidly increasing scholarly interest and investment in this research area. This steady and sharp increase across the years reflects the growing recognition of Federated Learning's potential for addressing critical privacy and security challenges in healthcare settings, as well as rising concerns related to data poisoning and adversarial attacks.

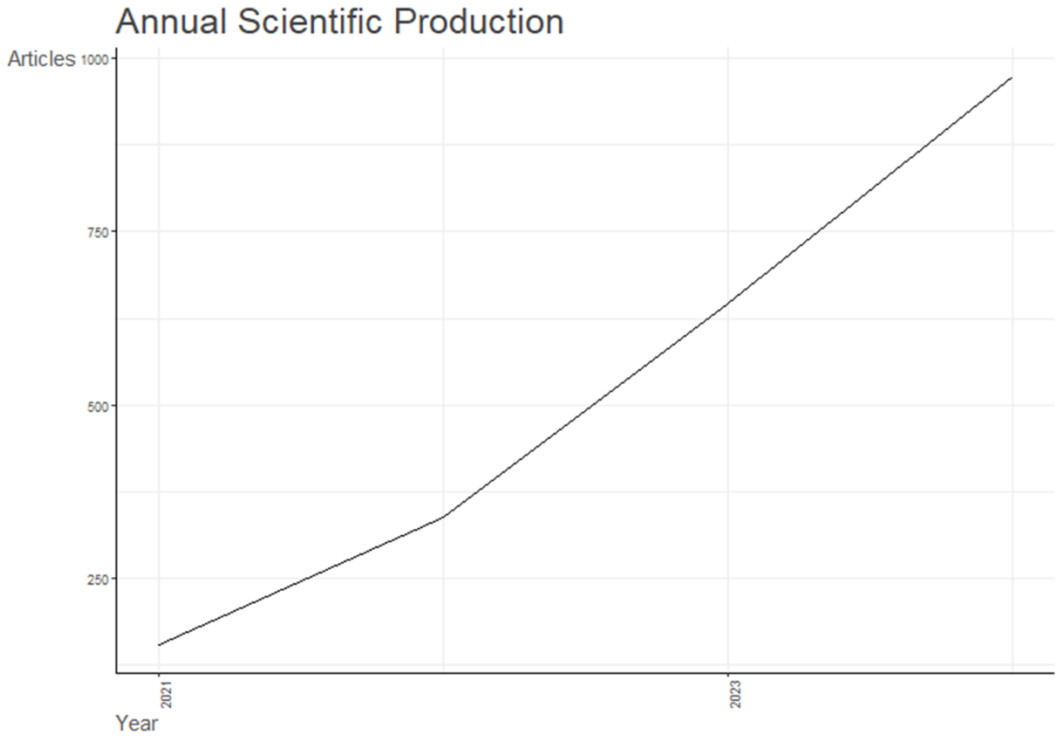

Fig. 1. Annual Scientific Production

5. MOST RELEVANT SOURCES

The analysis of publication sources (see Figure 2) revealed that "Lecture Notes in Computer Science" (including its subseries on Artificial Intelligence and Bioinformatics) emerged as the leading contributor with a substantial 117 articles. This journal significantly outpaced other publication outlets, emphasizing its pivotal role in disseminating research on federated learning in healthcare. "IEEE Access" followed, contributing notably with 59 articles, while "IEEE Journal of Biomedical and Health Informatics" presented 44 articles, highlighting the strong presence and influence of IEEE publications in this research domain.

Other relevant sources included "Lecture Notes in Networks and Systems" with 37 articles and "Communications in Computer and Information Science" with 35 articles. Additionally, specialized journals such as the "IEEE Internet of Things Journal" and "IEEE Transactions on Medical Imaging" demonstrated focused contributions, with 31 and 19 articles respectively, underscoring their specific engagement with healthcare technology and security concerns. Notably, "Sensors" and "Electronics (Switzerland)" each contributed around 19 to 20 articles, reflecting their interdisciplinary relevance to the broader technological applications and integration of federated learning. Overall, the diversity and depth of contributions from these sources underline the multidimensional research interest and growing importance of FL in addressing data security and privacy within healthcare.

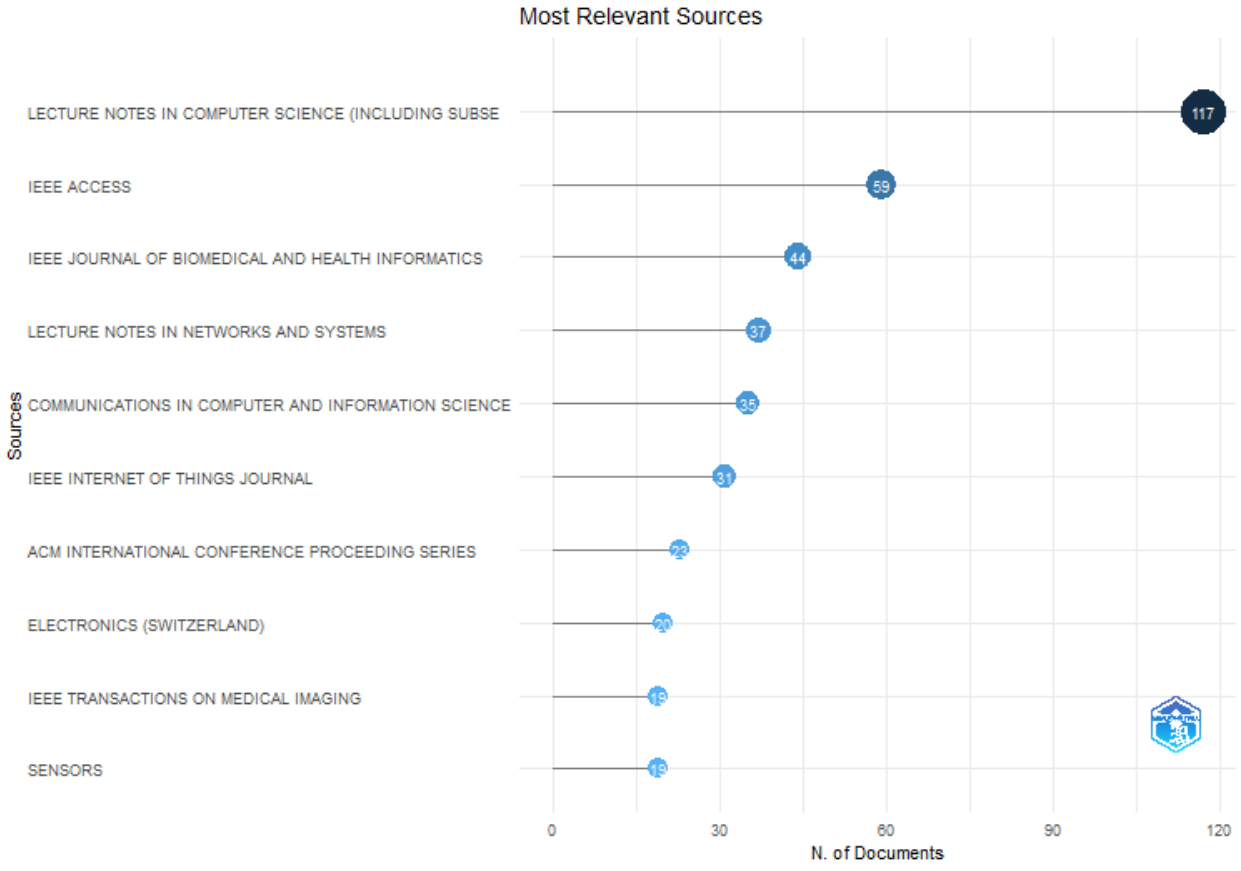

Fig. 2. Most Relevant Sources

6. MOST RELEVANT AFFILIATION

Analyzing affiliations contributing significantly to the research on Federated Learning in healthcare (see Figure 3), King Saud University and the Technical University of Munich emerged as the top contributors, each publishing 32 articles. Following closely, Chitkara University Institute of Engineering and Technology and the University of Electronic Science and Technology of China each contributed 30 articles, underscoring their substantial role in advancing research in this domain. Additionally, Huazhong University of Science and Technology contributed notably with 29 articles, highlighting the significant engagement of Chinese institutions.

Further notable affiliations include Chitkara University with 23 articles, reflecting a strong institutional focus in this field. The Lebanese American University, Shanghai Jiao Tong University, and Zhejiang University each contributed 22 articles, indicating robust international participation and collaboration. Finally, Imperial College London contributed 21 articles, illustrating Europe's active role in advancing federated learning research in healthcare. Collectively, these leading affiliations represent diverse geographical areas and reflect global interest and interdisciplinary collaboration in addressing data poisoning, privacy, and security within federated learning.

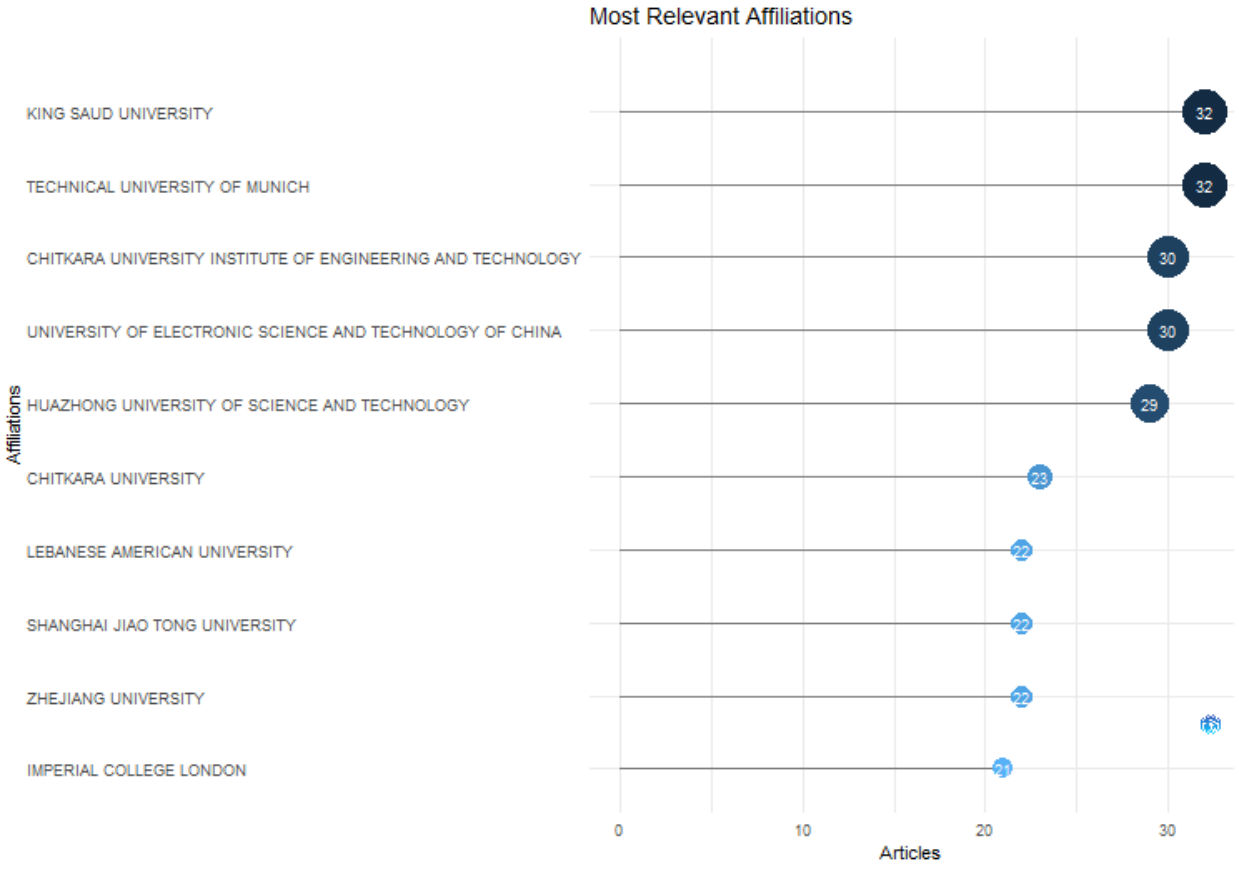

Fig. 3. Most Relevant Affiliation

7. AFFILIATIONS' PRODUCTION OVER TIME

Analyzing the production trends of top research affiliations over the years from 2021 to 2024 (see Figure 4), notable growth patterns emerged. The Technical University of Munich consistently maintained high productivity, increasing from 14 articles in 2021 to 32 articles by 2024, underscoring its sustained leadership in research activities related to Federated Learning in healthcare. Similarly, King Saud University demonstrated significant growth, rising sharply from a modest contribution of just 1 article in 2021 to an impressive output of 32 articles by 2024.

Chinese institutions such as Huazhong University of Science and Technology and the University of Electronic Science and Technology of China also exhibited consistent and notable growth. Huazhong University grew progressively from 6 articles in 2021 to 29 in 2024, while the University of Electronic Science and Technology of China increased remarkably from only 2 articles in 2021 to 30 articles in 2024, highlighting China's rising prominence in this research domain.

Interestingly, Chitkara University Institute of Engineering and Technology showed a delayed but rapid growth, contributing no publications initially in 2021 and 2022, but subsequently accelerating to 12 articles in 2023 and then sharply increasing to 30 articles in 2024. This trend emphasizes the expanding engagement and dynamic investment of specific institutions in Federated Learning research, particularly addressing data privacy, security, and adversarial threats within healthcare.

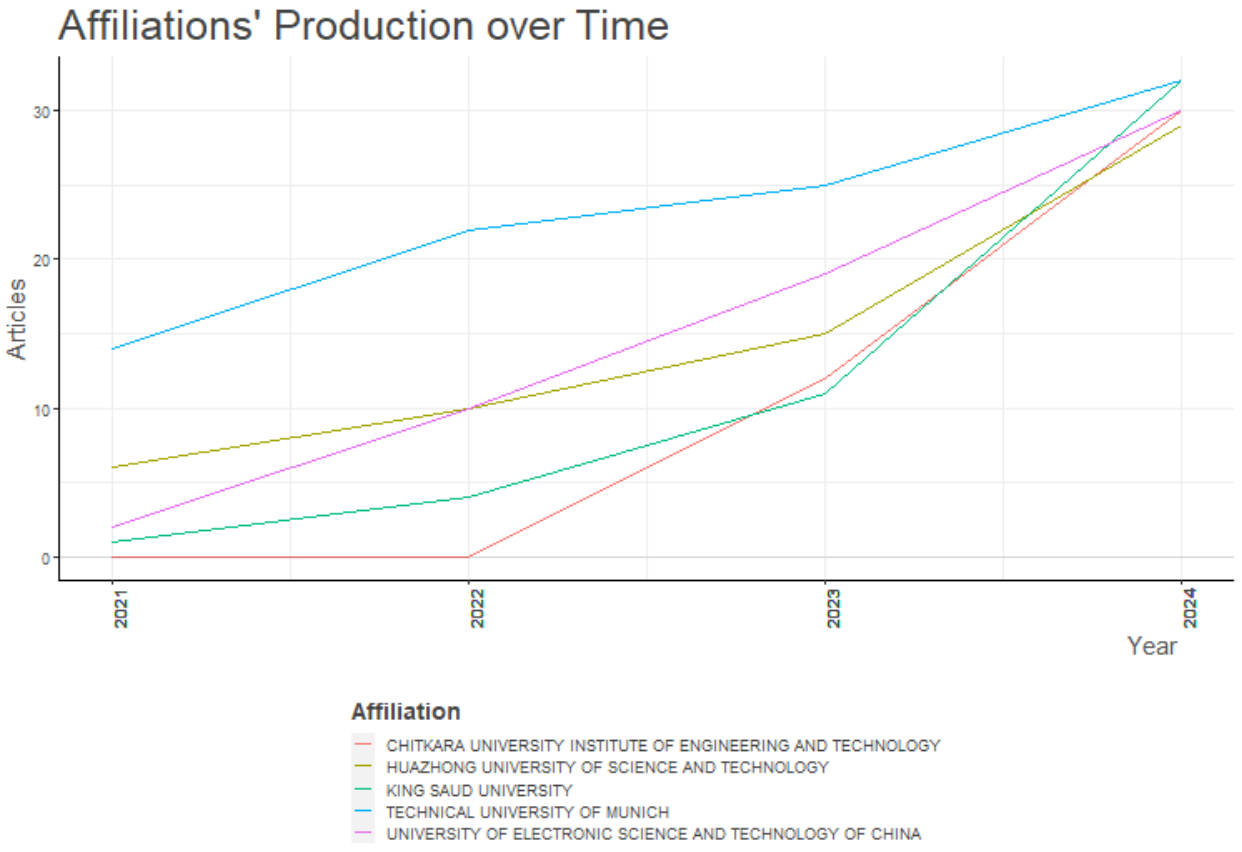

Fig. 4. Affiliations' Production Over Time

8. CORRESPONDING AUTHOR'S COUNTRIES

Analysis of corresponding authors' countries (see Figure 5) highlights a significant international presence in the research domain of Federated Learning in healthcare. China stands out prominently with 416 articles, showcasing a substantial mix of single-country publications (SCP: 263) and multiple-country collaborations (MCP: 153), reflecting a high international engagement indicated by an MCP ratio of 0.3678. India follows as the second-highest contributor with 249 articles, though demonstrating fewer international collaborations with an MCP ratio of 0.2811.

The USA contributes notably with 136 articles, primarily emphasizing single-country research (SCP: 103) but still maintaining a substantial collaborative engagement. Korea, despite having fewer total publications (87), shows a high international collaborative ratio (MCP: 34, MCP ratio: 0.3908), indicative of significant global cooperation.

European countries such as Germany, the United Kingdom, and Italy have also contributed actively, with Germany leading with 51 articles and a balanced international collaboration rate (MCP ratio: 0.4509). The UK and Canada, though contributing slightly fewer publications (46 and 38 respectively), exhibit notably high MCP ratios (0.5869 and 0.5789), reflecting significant international collaboration.

Countries like Saudi Arabia, Hong Kong, Pakistan, and Norway, although producing fewer articles, have impressively high MCP ratios (ranging from 0.6176 to 0.8125), demonstrating their strategic focus on international collaboration. Conversely, Japan's contributions lean towards domestic collaborations (MCP ratio: 0.2143), indicating a strong internal research network. These patterns collectively underscore the global interest and diverse collaborative dynamics shaping Federated Learning research in healthcare.

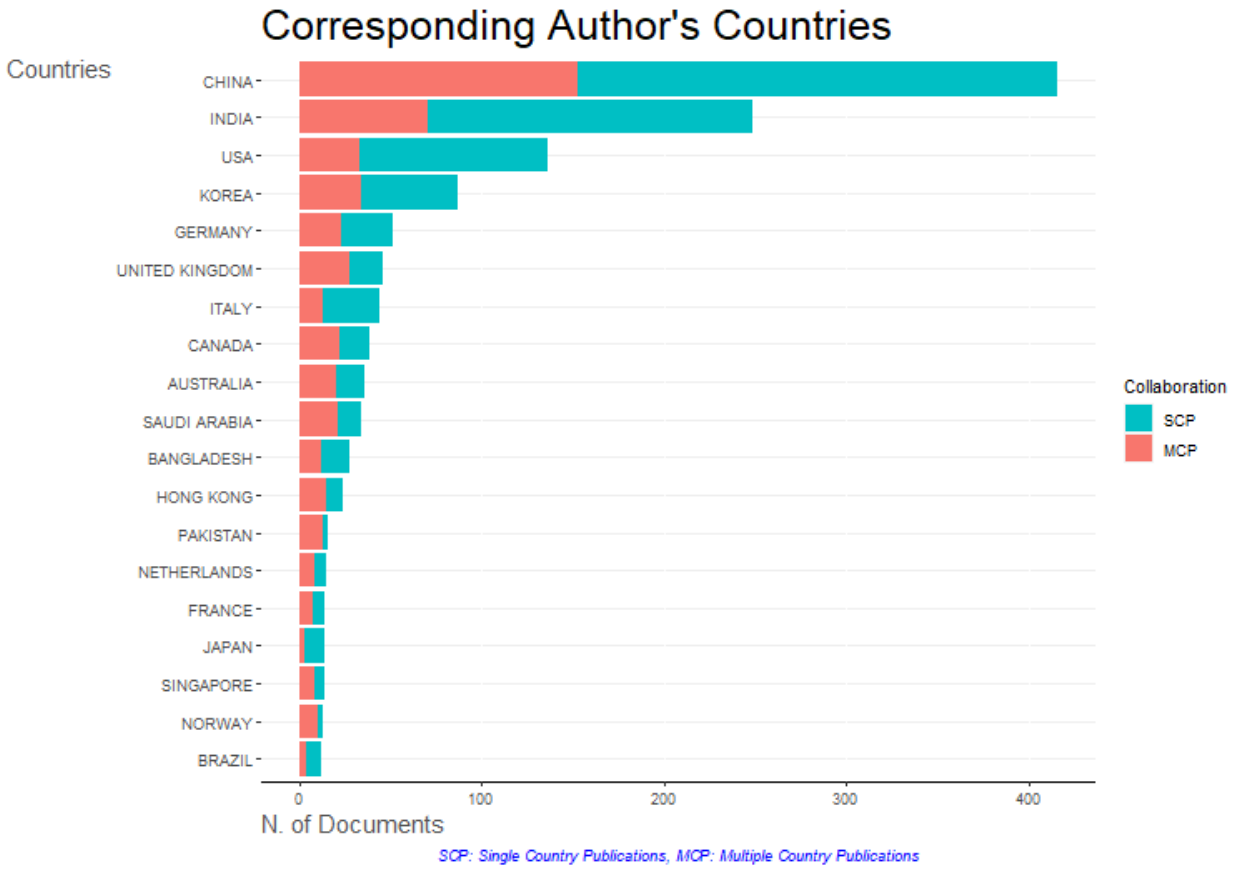

Fig. 5. Corresponding Author's Countries

9. COUNTRY SCIENTIFIC PRODUCTION

Figure 6 illustrates the frequency of scientific production across various countries, providing a snapshot of global research output. Notably, China leads with the highest frequency at 1376, significantly outpacing other nations and indicating a robust and prolific research ecosystem. India follows in second place with 1067, demonstrating a substantial scientific contribution, though still considerably lower than China. The United States holds the third position with 776, reflecting its strong and established scientific presence, albeit with a notable gap compared to the leading two countries.

A significant drop in frequency is observed as we move to Germany, which records 334, and the UK with 227, suggesting a tier of countries with moderately high scientific production. South Korea and Australia, with 213 and 191 respectively, maintain a consistent level of output, highlighting their growing contributions to global research. Saudi Arabia, with 172, and Italy, with 149, indicate a further decrease in frequency, yet still represent substantial scientific activity. Canada, with 148, mirrors Italy's output, suggesting a similar level of scientific production.

The data reveals a considerable decline in frequency as we move towards Pakistan and the Netherlands, with 108 and 102 respectively. Japan, with 86, Singapore, with 79, and France, with 76, form another group with relatively close frequencies, indicating consistent but lower output compared to the leading nations. Bangladesh and Iraq, with 72 and 69 respectively, show a further reduction, although they still maintain a notable presence in scientific production. Malaysia and the United Arab Emirates, with 64 and 61, demonstrate comparable levels of output.

Spain, with 55, and Greece, with 47, indicate a continued decrease in frequency, followed by Switzerland and Norway, both in the low 40s. Turkey and Portugal, with 36 and 34 respectively, show similar levels, while Egypt and Iran, both with 33, maintain a consistent but lower output. Brazil, Lebanon, and Sweden all register 29, indicating a similar level of scientific production among these countries. The data, when considered as a whole, highlights a clear disparity in scientific production, with a few nations dominating the field and a long tail of countries contributing at lower frequencies.

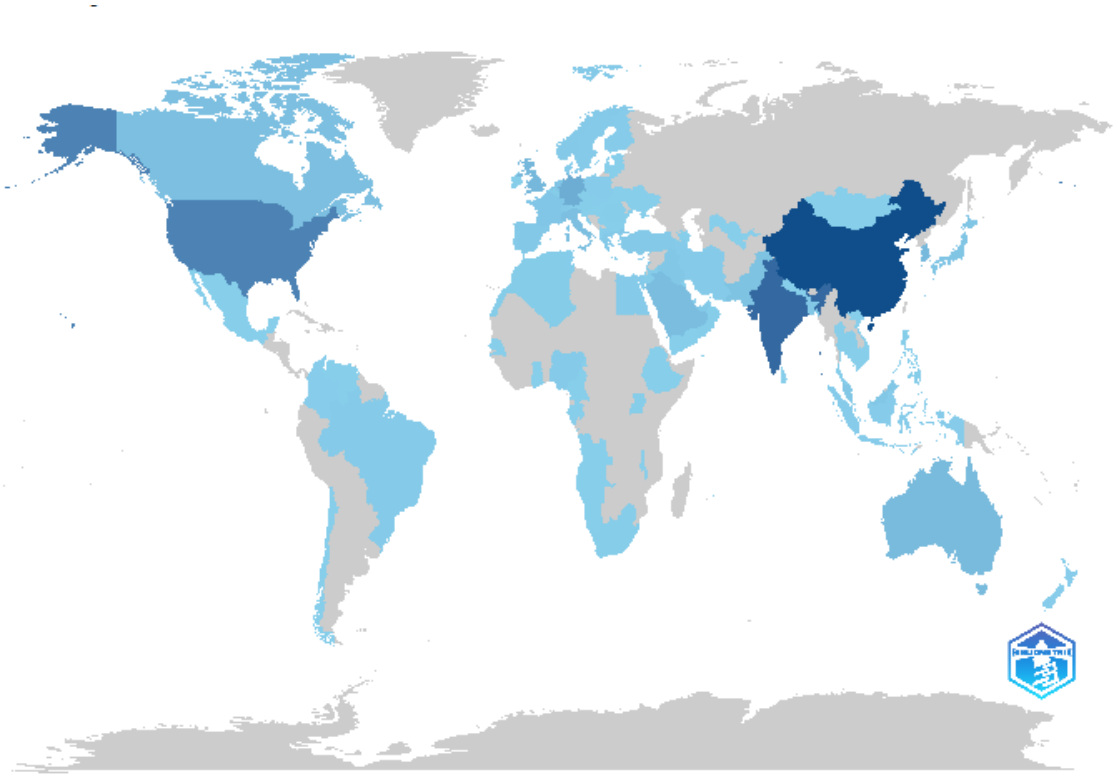

Fig. 6. country scientific production

10. WORD CLOUD

Figure 7 presents a word cloud illustrating the frequency of terms within a dataset, highlighting key areas of focus. The term "federated learning" stands out prominently with a frequency of 1245, indicating its significant prevalence and importance within the context of the data. This dominance suggests that federated learning is a central theme, likely driving much of the discourse or research represented by the dataset. Following "federated learning," "deep learning" and "learning systems" both appear with a frequency of 653, demonstrating their substantial relevance and correlation with the primary theme. The near-identical frequencies of these two terms suggest they are often used in conjunction or are closely related in the dataset.

"Data privacy" emerges as another critical term with a frequency of 476, underscoring the importance of privacy considerations in the context of the analyzed data. This is further reinforced by the presence of "privacy" (412), "machine learning" (447), and "machine-learning" (354), all of which indicate a strong focus on both the technical aspects of machine learning and the ethical considerations surrounding data privacy. The term "health care" with 344 and "medical imaging" with 333 highlight the specific application domains where these techniques are being applied, emphasizing the intersection of technology and healthcare. This is further supported by terms like "diagnosis" (314), "medical data" (125), and "diseases" (122), which collectively indicate a strong focus on medical applications.

"Privacy-preserving techniques" (323) and "privacy preserving" (254) further emphasize the importance of privacy considerations, suggesting that the development and application of such techniques are key areas of focus. Terms like "internet of things" (236), "adversarial machine learning" (234), "blockchain" (230), "block-chain" (206), and "artificial intelligence" (204) illustrate the broader technological landscape within which federated learning and data privacy are situated. "Differential privacy" (181), "sensitive data" (164), and "decentralised" (152) reinforce the focus on secure and privacy-centric approaches. Finally, terms like "algorithm" (135), "healthcare" (127), "hospital data processing" (124), and "edge computing" (121) highlight the technical and practical aspects of implementing these technologies in real-world scenarios. The word cloud, as a whole, underscores the interconnectedness of these terms, revealing a strong emphasis on federated learning, data privacy, and their applications in healthcare, within a broader context of advanced technologies.

Fig. 7. word cloud

11. CO-WORD NETWORK

The co-word network analysis (see Figure 8) highlighted critical thematic clusters within research on Federated Learning (FL) in healthcare. "Federated learning" emerged as the most central node, demonstrating the highest betweenness centrality (26.97), indicating its crucial role as an interdisciplinary bridge connecting various research themes. Closely related terms such as "learning systems," "deep learning," and "data privacy" also featured prominently, reflecting the significant intersections between machine learning methodologies and privacy preservation in healthcare data management. Other notable concepts included "machine-learning" and "medical imaging," signifying active research at the intersection of healthcare diagnostics and advanced algorithmic models. The presence of terms like "privacy-preserving techniques" and "privacy preserving" underscores the strong emphasis researchers place on developing methodologies to protect sensitive healthcare data against privacy breaches.

The analysis also revealed growing interest in integrating healthcare-specific terms such as "hospital data processing," "medical data," "diagnosis," and "diseases," suggesting that research increasingly targets practical medical applications. Emerging technologies such as "blockchain," "edge computing," and "internet of things" illustrate the ongoing exploration of new technological frameworks designed to enhance security and decentralization.

Moreover, the importance of "convolutional neural network" and "network security" signals attention toward developing robust and secure deep learning architectures. The nodes "classification," "image segmentation," and "electronic health record" suggest specific areas of application where federated learning techniques are making considerable impacts. This comprehensive co-word network thus encapsulates the intricate interplay of machine learning, security, privacy, and healthcare applications, delineating clear directions and focal areas within this dynamic field of research.

12. COLLABORATION WORLD MAP

The collaboration world map analysis (see Figure 9) illustrates extensive international collaboration within the domain of Federated Learning in healthcare. Australia appears notably active, with the highest frequency of collaboration observed with Saudi Arabia (7 occurrences), Bangladesh (6 occurrences), and Malaysia (6 occurrences). Canada's strongest collaborative link was with Lebanon, reflecting a high interaction frequency (9 occurrences), while also maintaining meaningful collaborations with Norway, Egypt, Poland, Pakistan, and Singapore.

Other notable collaborations include Bangladesh partnering frequently with Finland and Malaysia, demonstrating robust cross-regional research interactions. Smaller yet meaningful connections include partnerships involving Algeria with countries like Cameroon, Oman, and Yemen, indicating emerging research networks in the Middle East and Africa. The European research network displayed targeted but geographically diverse interactions, with Belgium collaborating with Denmark and Luxembourg, Austria interacting with Romania and Serbia, and Canada establishing broader networks across multiple continents.

Overall, these international collaborative patterns underscore the globalized nature of research efforts in federated learning, particularly regarding healthcare data security and privacy. Such international cooperation not only enriches the research landscape but also promotes the exchange of innovative methodologies and diverse perspectives essential for addressing complex global healthcare challenges.

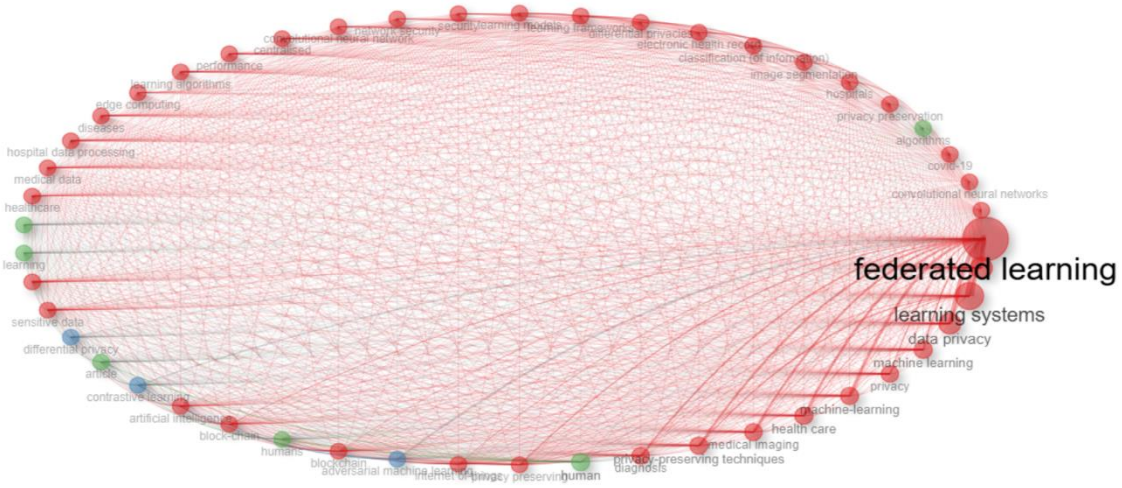

Fig. 8. Co-Word Network

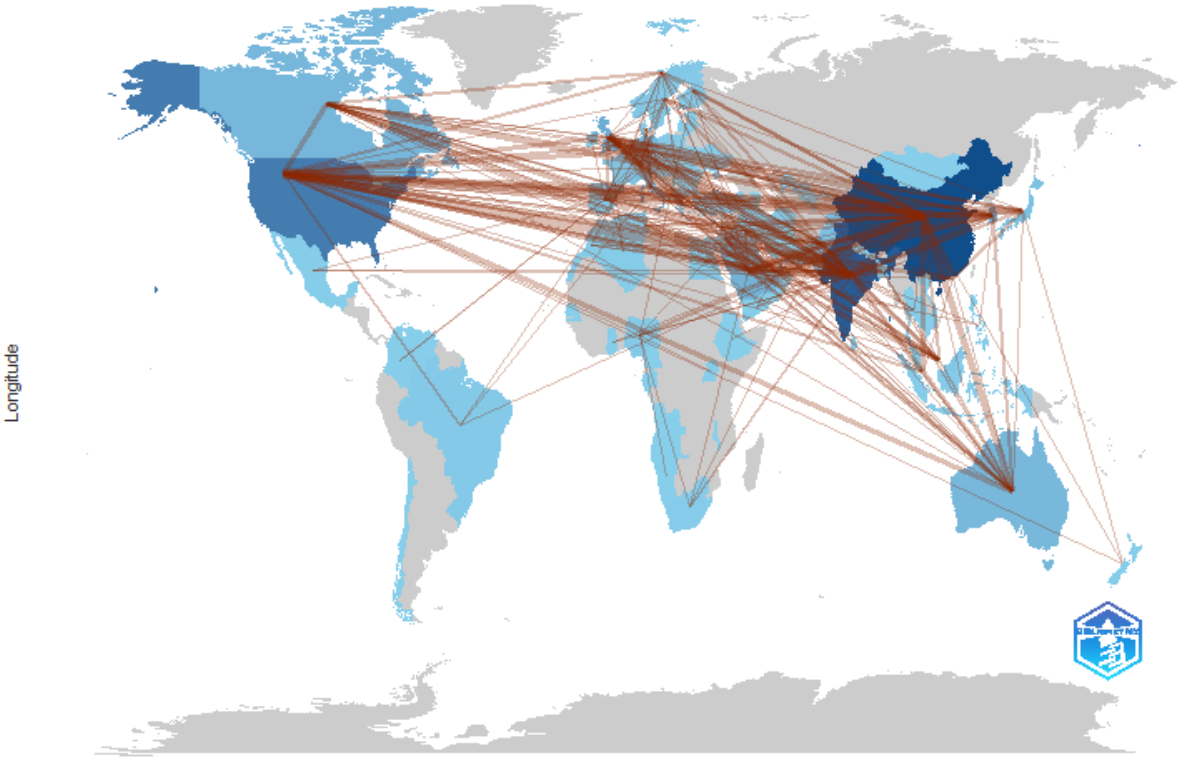

Fig. 9. Collaboration World Map

13. DISCUSSION

The bibliometric analysis of research on Federated Learning (FL) in healthcare from 2021 to 2024 reveals significant insights into the evolving landscape of this field. A primary observation from the co-word network analysis (see Figure 8) is that "federated learning" has established itself as the central node within the domain, featuring the highest betweenness centrality (26.97). This pivotal position demonstrates federated learning's fundamental role in connecting diverse research themes, particularly emphasizing machine learning methodologies, data privacy, and healthcare-specific applications.

The strong correlation of federated learning with "learning systems," "deep learning," and "data privacy" underscores an active research synergy between advanced computational methods and essential privacy-preserving strategies. This interplay highlights the growing recognition within the scientific community of the need to reconcile technological advancements with stringent privacy requirements inherent to medical data. Moreover, the notable presence of terms such as "machine-learning" and "medical imaging" indicates that federated learning is frequently positioned at the nexus of technology-driven medical diagnostics and data security.

The thematic prominence of "privacy-preserving techniques" and "privacy preserving" further emphasizes that privacy considerations are not merely adjunct concerns but integral to the practical implementation of federated learning in healthcare environments. This persistent focus reflects an acute awareness of potential vulnerabilities, particularly data poisoning and adversarial attacks, which represent significant threats to data integrity and patient privacy. These findings align closely with the prevalent concerns identified in the initial research query targeting explicitly "data poisoning," "adversarial attack," "privacy," and "security."

Healthcare-specific terminology like "hospital data processing," "medical data," "diagnosis," and "diseases" demonstrates an increasing practical orientation in research, suggesting a shift toward real-world clinical applications of federated learning. These terms collectively highlight researchers' interest in solving tangible, domain-specific challenges, illustrating the evolving maturity of federated learning from theoretical exploration to actionable implementation. Furthermore, the emergent technological terms "blockchain," "edge computing," and "internet of things" (IoT) suggest a deliberate exploration into hybrid solutions aimed at reinforcing security, decentralization, and operational efficiency in healthcare data handling.

Interestingly, terms associated with deep neural architectures, notably "convolutional neural network" (CNN) and "network security," signify a robust commitment to developing computationally sophisticated yet secure learning frameworks. The presence of CNN-related research indicates a sustained investment in leveraging high-performance algorithms for precise medical diagnostics, while concurrently ensuring data privacy through secure FL frameworks. Additionally, concepts such as "classification," "image segmentation," and "electronic health record" (EHR) align closely with practical healthcare applications, further suggesting an explicit alignment of FL research with healthcare service enhancement and medical innovation.

The international dimension of federated learning research is clearly illustrated through the collaboration world map analysis (see Figure 9). Australia emerges as notably proactive, evidenced by frequent collaborative engagements with countries across diverse regions, including Saudi Arabia, Bangladesh, Malaysia, and Japan. This broad-based collaboration reflects Australia's strategic focus on enhancing research impact through international partnerships, particularly in regions where federated learning can significantly benefit healthcare systems facing privacy and security challenges.

Canada's notably high-frequency collaborative relationship with Lebanon underscores unique bilateral interactions, reflecting potential research synergies possibly driven by shared research agendas or complementary expertise. Canada's other significant collaborations with countries such as Egypt, Norway, Pakistan, Poland, and Singapore suggest robust North-South and transcontinental research linkages, indicative of a strategic and geographically diverse partnership approach. Such cross-regional partnerships are critical in federated learning, given its inherently decentralized and collaborative nature.

Furthermore, the presence of significant yet regionally targeted collaborations, such as Bangladesh with Finland and Malaysia, indicates focused efforts toward building specialized research capabilities and knowledge exchange. These targeted interactions can stimulate localized innovations, supporting healthcare sectors that are particularly vulnerable to privacy and security threats.

Emerging collaborative relationships, such as those between Algeria and Middle Eastern or African nations (Cameroon, Oman, and Yemen), highlight new and potentially strategic regional research networks. These nascent interactions suggest opportunities for future growth and innovation in federated learning applications, particularly within regions experiencing rapid digital health adoption amidst significant infrastructural challenges.

European collaborations, although smaller in individual frequency, exhibit considerable geographical diversity, with Belgium engaging Denmark and Luxembourg, and Austria collaborating with Eastern European countries like Romania

and Serbia. This diversity highlights Europe's complex yet interconnected research landscape, reflecting a cohesive approach to federated learning issues of data privacy, security, and adversarial threats in healthcare.

Comparative insights from corresponding author countries (see Figure 5) further contextualize global federated learning research dynamics. China's dominant position, producing 416 articles, alongside a substantial percentage of international collaborations (MCP ratio: 0.3678), indicates significant research leadership complemented by active international engagement. India's robust domestic research output (249 articles), albeit with a slightly lower international collaboration rate (MCP ratio: 0.2811), suggests a strong internal research infrastructure paired with gradually increasing global outreach. Contrastingly, countries like Germany and the United Kingdom, despite producing fewer total articles, demonstrate relatively high international collaboration rates (MCP ratios: 0.4509 and 0.5869, respectively). This indicates that European countries are strategically leveraging international research cooperation to amplify their scientific impact and address complex interdisciplinary challenges in federated learning.

Additionally, countries like Saudi Arabia, Hong Kong, Pakistan, and Norway, despite lower absolute publication counts, exhibit impressively high MCP ratios (ranging from 0.6176 to 0.8125), emphasizing a strategic emphasis on international cooperation. This approach suggests these nations effectively utilize global research networks to overcome limitations in local research capacities, enriching their research environments with global expertise and diverse insights.

Finally, examining scientific production at the country level (Figure 6), China and India notably dominate, with outputs significantly exceeding those of other nations. The substantial gap between these two countries and others like the USA, Germany, and the UK underscores distinct regional research leadership dynamics. This pattern reflects intense research activity and governmental or institutional prioritization of federated learning technologies in these leading countries, aligning closely with their increasing investment in healthcare digitization and cybersecurity.

In conclusion, the discussion synthesized from the bibliometric analysis offers an extensive view of federated learning research within healthcare, emphasizing its global nature, interdisciplinary, and practical orientation. The insights gained from thematic clusters, international collaborations, and country-specific production dynamics collectively suggest federated learning's critical role in future healthcare solutions, particularly in addressing pervasive concerns surrounding data privacy, security, and adversarial threats. These findings highlight essential avenues for future research, emphasizing the need for sustained international collaboration, strategic partnerships, and targeted interdisciplinary engagements to further advance this vital area of research.

14. CONCLUSION

This bibliometric analysis has systematically explored the current landscape of federated learning (FL) research within the healthcare domain, emphasizing key aspects such as data privacy, security threats, and adversarial attacks. The findings reveal rapidly increasing scholarly interest, particularly in privacy-preserving methods, healthcare-specific applications, and emerging technologies like blockchain and edge computing. Moreover, robust international collaboration and significant contributions from leading institutions and countries demonstrate a collective effort toward addressing critical security and privacy concerns within FL applications.

However, the study also identified critical limitations, primarily stemming from incomplete bibliographic metadata, highlighting the necessity for improved data collection and management practices in scholarly databases. Addressing these limitations is essential for accurately monitoring research trends and facilitating future investigations.

Overall, this research provides valuable insights for researchers, practitioners, and policymakers, assisting them in better understanding existing challenges and opportunities within federated learning for healthcare. Future research directions include advancing comprehensive security frameworks, enhancing collaborative research efforts, and continuously updating approaches to address emerging threats and technological advancements in the healthcare sector.

Funding:

This research was not funded by any institution, foundation, or commercial entity. All expenses related to the study were managed by the authors.

Conflicts of Interest:

The authors declare that there are no conflicts of interest to disclose.

Acknowledgment:

The authors wish to acknowledge their institutions for their instrumental support and encouragement throughout the duration of this project.

References

[1] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, “Federated learning for internet of things: A comprehensive survey,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1622–1658, 2021.

[2] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, “Federated learning: A survey on enabling technologies, protocols, and applications,” *IEEE Access*, vol. 8, pp. 140699–140725, 2020.
[3] T. Zhang, L. Gao, C. He, M. Zhang, B. Krishnamachari, and A. S. Avestimehr, “Federated learning for the internet of things: Applications, challenges, and opportunities,” *IEEE Internet Things Mag.*, vol. 5, no. 1, pp. 24–29, 2022.
[4] R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari, and B. Eskofier, “Federated learning for healthcare: Systematic review and architecture proposal,” *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 4, pp. 1–23, 2022.
[5] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, “Federated learning for healthcare informatics,” *J. Healthc. Inform. Res.*, vol. 5, pp. 1–19, 2021.
[6] R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari and B. Eskofier, "Federated learning for healthcare: Systematic review and architecture proposal," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 13, no. 4, pp. 1–23, 2022.
[7] J. Li *et al*., “A federated learning-based privacy-preserving smart healthcare system,” *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, 2021.
[8] Lakhan *et al*., “Federated-learning based privacy preservation and fraud-enabled blockchain IoMT system for healthcare,” *IEEE J. Biomed. Health Inform.*, vol. 27, no. 2, pp. 664–672, 2022.
[9] K. Kaleel, A. Assad, and M. Fyadh, Trans., “Quranic studies evolution: A bibliometric analysis from 1880 to 2023,” *Mesopotamian J. Quran Stud.*, 2024, pp. 1–12. [Online]. Available: https://doi.org/10.58496/MJQS/2024/001
[10] Benseng and S. Tam, Trans., “Mapping the evolution of Arabic language research: A bibliometric approach,” *Mesopotamian J. Arabic Lang. Stud.*, 2024, pp. 1–15. [Online]. Available: https://doi.org/10.58496/MJALS/2024/001
[11] Yazdinejad, A. Dehghantanha, H. Karimipour, G. Srivastava, and R. M. Parizi, “A robust privacy-preserving federated learning model against model poisoning attacks,” *IEEE Trans. Inf. Forensics Secur.*, 2024.
[12] S. Pati *et al*., “Privacy preservation for federated learning in health care,” *Patterns*, vol. 5, no. 7, 2024.
[13] Murmu, P. Kumar, N. R. Moparthi, S. Namasudra, and P. Lorenz, “Reliable federated learning with GAN model for robust and resilient future healthcare system,” *IEEE Trans. Netw. Serv. Manag.*, 2024.
[14] E. Darzi, F. Dubost, N. M. Sijtsema, and P. M. van Ooijen, “Exploring adversarial attacks in federated learning for medical imaging,” *IEEE Trans. Ind. Informat.*, 2024.