*Çanakkale Onsekiz Mart University*

*Journal of Advanced Research in Natural and Applied Sciences*

dergipark.org.tr/tr/pub/jarnas

Open Access



e-ISSN: 2757-5195 | 2024, Vol. 10, Issue 4, Pages: 875-898 | doi.org/10.28979/jarnas.1555351

# A Bibliometric Analysis on Federated Learning

Ömer Algorabi[1], Yusuf Sait Türkan[2], Mesut Ulu[3], Ersin Namlı[4]

[1,2,4]Department of Industrial Engineering, Istanbul University, Cerrahpaşa, İstanbul, Türkiye
[3]Department of Occupational Health and Safety, Bandırma Onyedi Eylül University, Balıkesir, Türkiye

**Article Info**

*Received*: 24 Sep 2024

*Accepted*: 22 Dec 2024

*Published*: 31 Dec 2024

Research Article

**Abstract** − With the rapid advancement of technology and growing concerns about data privacy, federated learning (FL) has attracted considerable attention from the scientific community. The emergence of FL as a novel machine-learning approach and the volume of relevant papers and studies now call for a thorough investigation of FL. In the present research, an analysis was conducted on 3107 articles about federated learning exported from the Web of Science (WoS). The paper performs a bibliometric analysis to examine the productivity, citations, and bibliographic matching of significant authors, universities/institutions, and countries. The evolution of research material on federated learning over time was analyzed in the research. The study also provides comprehensive analysis by examining the most frequently used terms in the articles and attempting to identify trending areas of study with federated learning. This paper offers primary information on FL for readers worldwide and a comprehensive and accurate analysis of potential contributors.

**Keywords –** *Federated learning, bibliometric analysis, network analysis*

## 1. Introduction

Federated Learning (FL) is an emerging machine learning approach that addresses the problem of data aggregation and, at the same time, ensures data privacy. This approach integrates the server and clients (organizations, mobile devices, etc.) to develop a decentralized machine-learning operation. This technology was first introduced by Google in 2016. In the first study, the prediction of data input on devices was developed in the case where data is stored on mobile devices [1]. After the first study on FL, the number of studies in this field in literature is rapidly increasing, and its application areas are expanding. In this context, it has become essential to examine the course of the different studies in this field and the conceptual dimensions and applications of FL. Federated learning is, in a sense, a cryptic distributed machine learning scheme. In the system, users' data is not disclosed. The binding of data is achieved by combining local parameter information. Thus, data privacy is ensured, and a common machine-learning model can be developed between clients [2]. Considering the General Data Protection Regulation (GDPR), the federated learning framework has evolved rapidly in recent years. With the rapid increase in the use of artificial intelligence, the federated learning approach has been used in many applications, such as smartphones [3], IoT [4-6], healthcare [7-9], advertising [10], autonomous vehicles [11-12], energy forecasting [13-15], fraud detection [16], and insurance [17].

FL can be classified into Horizontal Federated Learning (HFL), Vertical Federated Learning (VFL), and Transfer Federated Learning (TFL) based on data partitioning. In HFL, the clients' data have the same features, but the sample space is different. In VFL, the IDs of the data come from the same sample, but the features are

[1]omer.algorabi@iuc.edu.tr; [2]ysturkan@istanbul.edu.tr; [3]mesutulu@beykent.edu.tr (Corresponding Author); [4]enamli@iuc.edu.tr

distinct. In the FTL scenario, the sample and attributes of the data differ mainly, while only a small part is similar [18]. There have been remarkable literature studies of these three types in recent years. Li et al. [19] developed a horizontal federated learning model for demand forecasting of e-commerce enterprises. Another study presents a system based on horizontal federated learning for automatically detecting IoT devices [20]. Wang et al. [21] proposed a retail supply chain demand forecasting model with a vertical federated learning framework. A data privacy-based system was developed using social media, e-commerce platforms, and Retailers' data. In a study conducted in 2023, vertical federated learning was used in energy load forecasting. A smart model is presented by incorporating three data sets in electricity demand, electricity generation, and energy price forecasting studies [22]. A federated transfer learning technique study presented a predictive model for cross-domain applications in smart manufacturing systems. They used two datasets of multiple images of pedestrians walking [23]. Chen et al. [24] proposed a federated transfer learning framework for Parkinson's disease diagnosis with smart wearable healthcare.

The privacy mechanism differs in the studies depending on the federated learning framework. The privacy mechanism is related to the aggregation of client data. Ye et al. [25] developed a model aggregation method for federated learning-based image processing. In this method, the clients' parameters in local learning are aggregated on the server. In a 2021 study for the communication of federated learning-based 6G-enabled vehicular networks, the parameter information in different layers of neural networks is aggregated by weighted averaging [26]. Another privacy mechanism used in studies is homomorphic encryption. Ou et al. [27] presented a vertical federated learning model with homomorphic encryption for risk management. Encryption is used to exchange model information. In the study conducted in 2022, a federated learning approach was used to classify medical images in an IoT-based healthcare system, and data privacy was ensured by homomorphic encryption [28]. Zhao et al. [29] presented a federated learning framework for intelligent traffic management using differential privacy. The differential privacy technique adds random noise to the aggregated data and prevents attacks on vulnerable data. Another study analyzed histopathological medical images using federated learning and differential privacy. A method was proposed that interrupts gradient updates and inserts noise in histopathology [30]. Jia et al. [31] presented a blockchain-enabled federated learning approach in the IIoT. Data protection is achieved through homomorphic encryption and differential privacy-based k-means clustering. Another privacy mechanism that is rarely encountered in studies is zero-knowledge systems. In these systems, clients only learn about the output. While these systems provide suitable security, their computational and communication costs are expensive [32]. Chen et al. [33] proposed a zero-knowledge clustering approach with federated learning against attacks on IoT data.

Researchers have developed several methods for data training of federated learning systems. In the literature, these methods can be categorized into three groups: statistical methods, methods based on decision trees, and methods based on neural networks. Gogineni et al. [34] developed a kernel regression model based on Fourier features for training local models. Wei et al. [35] presented a two-part logistic regression method in the learning process of a vertical federated learning system. In another vertical federated learning architecture, a method based on ridge regression is proposed for bi- and multi-party scenarios [36]. When the studies in which decision trees are preferred are examined, Yamamoto et al. [37] presented a gradient-boosting decision tree model based on federated learning for the privacy-based learning process using publicly available datasets. Studies such as financial risk prediction [38], intrusion detection [39], cancer disease prediction [40], etc., have been performed using the random forest method in federated learning applications. In a study predicting hospitalizations due to cardiac events, a support vector machine methodology was developed within the framework of federate learning [41]. Due to the nature of federated learning, neural network methods have been widely used by researchers to address privacy concerns. Vaid et al. [42] used the multilayer perceptron method for local training in a federated learning framework to predict mortality in hospitalized COVID-19 patients. A study using MR images performed local training with deep learning models [43]. LSTM, one of the deep learning methods, is another method used in the federated learning system [44–45]. Metaheuristic algorithms have been widely used for parameter optimization. Genetic algorithms [46], simulated annealing [47], particle swarm optimization [48], etc., have been used for parameter optimization of local training in federated learning systems. Cui et al. [49] utilized federated learning technology and proposed an adaptive neuro-fuzzy inference system (ANFIS) optimized with gray wolf optimization to predict the reliability of groundwater levels. The literature also includes fuzzy models in the federated learning system [50–52].

In the short history of federated learning, there have been several reviews of the approach. KhoKhar et al. [53] reviewed image processing using federated learning. Reviews have been examined on the application of Federated Learning intrusion detection [54], resource optimization [55], healthcare [56], cyberspace security [57], electric vehicle transportation [58], 6G [59], etc. Javed et al. [60] surveyed vehicle networks and intelligent transportation systems using integrated federated learning and blockchain technologies. A study in 2023 presented a comprehensive survey of attacks and defenses against federated learning [61].

Our study aims to provide a comprehensive bibliometric analysis of federated learning publications. Our study was conducted to fill a research gap in the field of FL. In 2018, there were only two studies on FL, and if we look at the years 2022 and 2023, the fact that more than a thousand studies are proposed annually has led to the need for a comprehensive literature review in this field. As there is no bibliometric study on FL from a holistic perspective, this comprehensive study on FL from a broad perspective is expected to significantly contribute to the literature. The publications discussed in this study were obtained from the WoS database, and then a comprehensive bibliometric analysis was carried out. The content of the study includes (1) publication types, number of publications, citations, and timeline analysis; (2) analysis of prolific authors, countries, organizations, document sources, and funding agencies; (3) collaborative networks using R; (4) keyword identification.

Bibliometric analysis is a popular and meticulous approach for surveying and analyzing large volumes of scientific information [62–63]. This method allows us to statistically analyze the developments in a given field and make predictions in this area. In recent years, bibliometric analysis has gained significant attention. Li et al. [64] analyzed deep learning studies between 2007 and 2019 using VOS Viewer and Cite Space. The study analyzed annual publications, citations, most productive authors, institutions, countries, journals, and publications. In another study where the same analyses were carried out, operations research and management science publications in the Web of Science database were examined [65]. Ezugwu et al. [66] presented an extensive study by performing taxonomic classification and bibliometric analysis of metaheuristic algorithms. Shukla et al. [67] analyzed engineering applications in artificial intelligence from 1998 to 2018, indexed in the Web of Science and Scopus. Big data analytics and machine learning studies were analyzed in another bibliometric analysis of artificial intelligence research, 2006-2020. The articles from Scopus were analyzed by grouping them into clusters according to their fields of study [68]. Yu et al. [69] examined the studies published in the Fuzzy Optimization and Decision-Making journal between 2002 and 2017. In addition to analyzing the most productive publications, authors, countries, and organizations, the co-citation networks of cited authors, sources, and references were also revealed. Zhang et al. [70] conducted a topic change analysis of the Knowledge-Based Systems journal using bibliometric analysis. The study aims to predict future trends.

In this paper, after the introduction and the extensive literature review, Section 2 describes the data and methodology of the bibliometric analysis. Section 3 presents a comprehensive bibliometric analysis of FL. The bibliometric analysis examined the most productive authors, countries, organizations, and journals. Collaborative studies were revealed through network diagrams, and analyses were conducted on frequently used terms. Finally, Section 4 discusses future research directions and concludes the study.

## 2. Bibliometric Analysis

### 2.1. Data Source and Methodology

Web of Science is a database that lists the impact of scientific journals published in various disciplines, the number of citations received by published articles, authors' articles, and bibliographies of articles. WoS, also one of the most widely used databases for bibliometric analyses, indexes high-quality publications [69]. Most review studies on publication, author, time, and other titles are done with the WoS database. The Web of Science Core Collection includes the Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (AHCI), Emerging Sources Citation Index (ESCI), Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index – Social Sciences & Humanities, Book Citation Index – Science, Book Citation Index – Social Sciences & Humanities citation indexes and covers titles from 1900 to the present day [71].

In our study, after reviewing planning, data collection, data analysis, and data visualization steps were carried out. As shown in Figure 1, opinions and suggestions for future research on federated learning were presented after all the steps. Before the data collection phase, the search topic was federated learning. Then, 5,740 publications were searched in the WoS repository. The earliest study in this field was published in 2017. The bibliometric analysis filtered the document type to articles only, and 3,107 searched publications were considered. Since our research aims to examine the studies on federated learning from a broad perspective, no special restrictions have been set for the data to be utilized. On 15 August 2023, all information on the articles included in the bibliometric analysis was transferred from WoS. The bibliometric analysis was initiated in August 2023. Accordingly, publications from 2024 could not be incorporated into the study. Further studies are planned, including an analysis of the current literature.

The exported data contains many attributes such as author, country, author institution, date, title, abstract, keywords, etc. In the bibliometric analysis, tables and figures are presented using total publications (TP), total citations (TC), citations per year, citations per publication (TC/TP), and h index performance indicators. Cooperation networks depict collaborations in the studies. Topics and concepts were analyzed with time graphs, thematic maps, and structure maps.

## 2.2. Data Source and Methodology

Using the WoS database, 5740 publications were found when the subject title was searched as federated learning. The publication types of these studies are given in Figure 2. Most studies were articles (3,107 publications) and proceeding papers (2460 publications). These two document types account for 97% of total publications. In early access and review article types, 204 and 129 studies were reported, respectively. It has been determined that there are insufficient studies on the publication types of correction, book chapters, letters, and book reviews. The number of publications in these document types is reported to be less than 10.

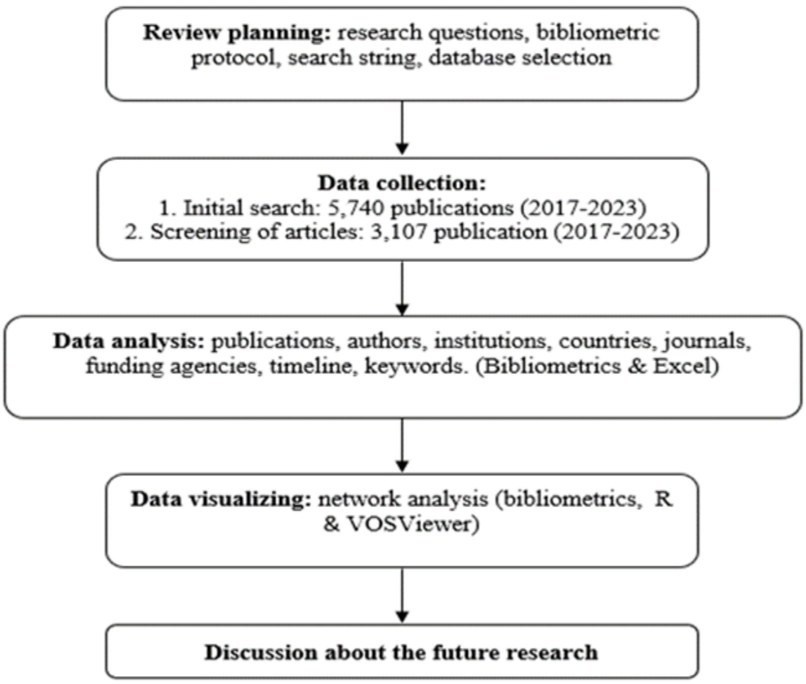



**Figure 1.** Research design

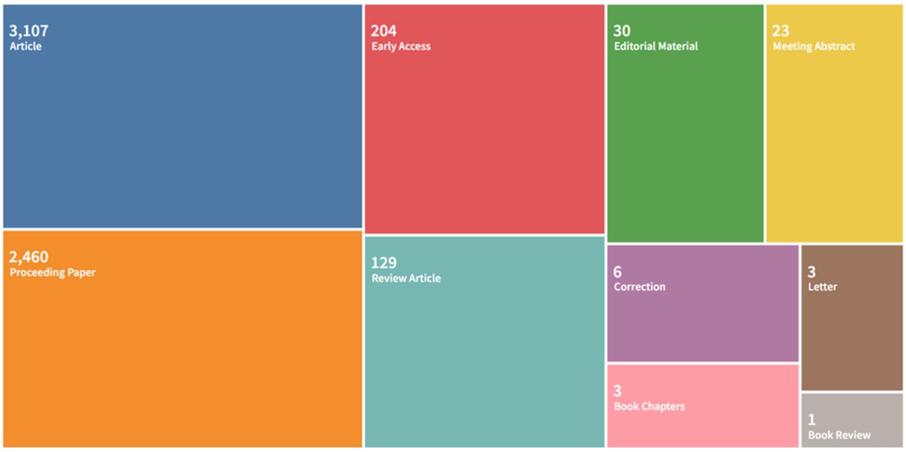



**Figure 2.** Document types of the publications

The categorization of article studies according to WoS is shown in Figure 3. It was found that article publications were primarily included in the study categories of engineering, electrical engineering, computer science, information systems, and telecommunications. The number of federated learning publications in these categories was more than 1000. After these categories, the fields with the highest number of articles are Computer Science Artificial Intelligence (381 articles) and Computer Science Theoretical Methods (355 articles). In particular, federated learning is known to be highly relevant to the disciplines of electrical and electronic engineering and computer science. The classification of federated learning studies according to research areas is given in Figure 4. Computer science and engineering fields account for 70% of the studies. Mathematics, physics, chemistry, materials science, and other disciplines contributed less than 10% to federated learning studies.

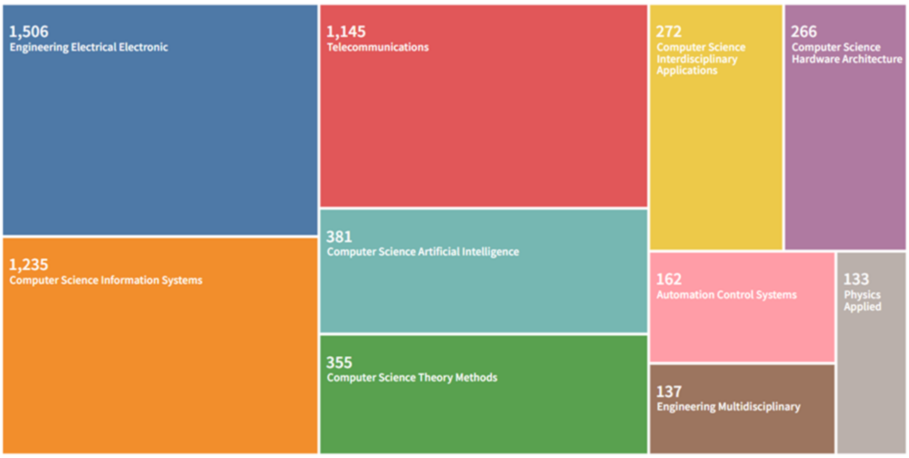



**Figure 3.** Top 10 WOS categories of the publications

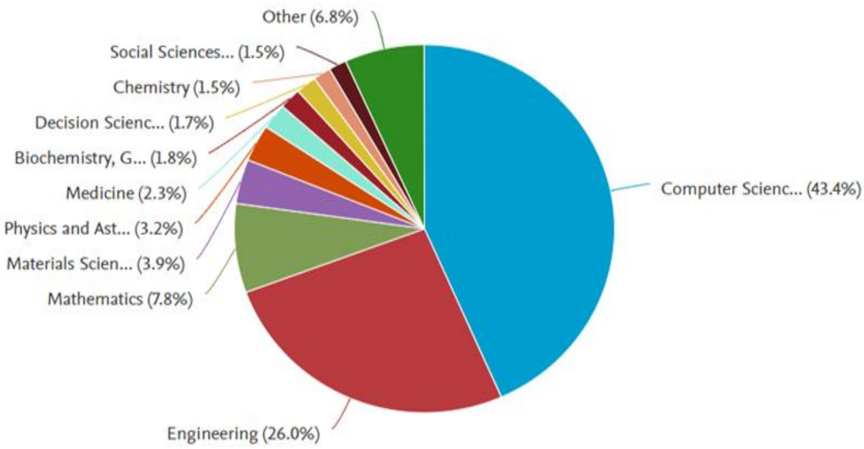



**Figure 4.** Research areas of the publications

According to the data obtained from the WoS database, the annual number of federated learning studies is illustrated in Figure 5. According to the graph, the first publication belongs to 2017. In 2019, studies in this field were limited to 12 articles. In 2020 and beyond, federated learning studies gained momentum, and the number of publications gradually increased. In 2020, the number of articles exceeded 100 for the first time, and in 2022, the number of articles exceeded 1,000. The number of publications in 2022 increased 2.13 times compared to 2021. The number of publications realized in 2023 (until August) is 1,080.

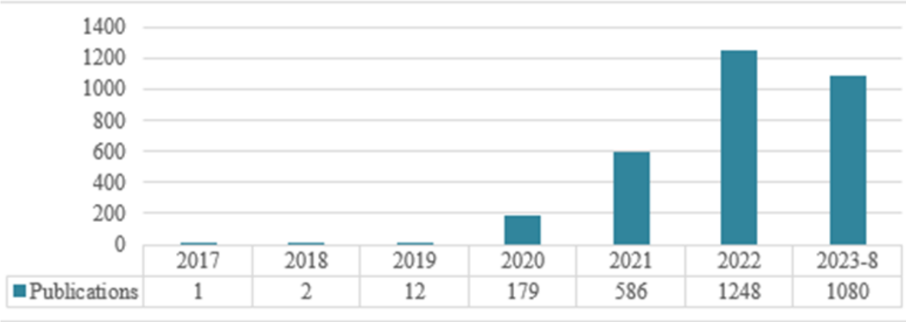



**Figure 5.** Annual publications

Figure 6 shows the annual number of citations and annual citations per publication. According to article studies conducted between 2017 and 2023-8, it was stated that there was an average of 8.83 citations per publication. It has been determined that 2020 publications received the most citations. This year's studies were cited 11276 times, and the average number of citations per article was 63. The highest TC/TP ratio was in 2019, with 482 citations per article. According to these results, it has been determined that the studies conducted in 2019 and 2020 are valuable, worth citing, and essential sources of current studies.

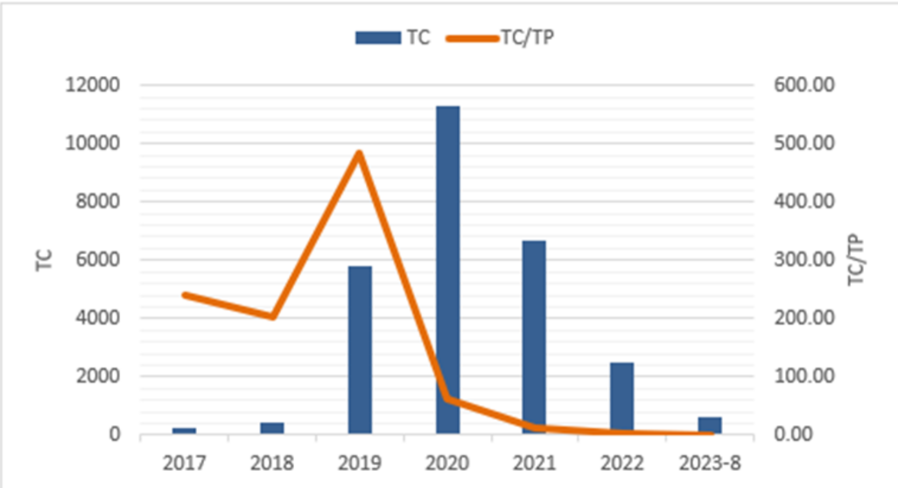



**Figure 6.** Citation trend analysis

The top 10 most influential publications, according to the highest number of citations, are given in Table 1. Yang et al. [72] is the most influential article with 3991 citations. All 10 articles have been cited more than 400 times, while the top 3 have been cited more than 800 times. On average, nine articles in the top 10 are cited 100 or more times per year. When the articles are analyzed, it is seen that all of them have three or more authors. Institutions in China conducted a Yang et al. [72] study. Li et al. [64] article was found to be prepared by USA institutions. Except for the 8th-ranked publication, it was determined that the articles ranked 3-10 were realized with the countries' cooperation.

### 2.2.1. Authorship and Institution Analysis

Since 2019, the authors have carried out many studies with federated learning. The most prolific authors in this field are given in Table 2. Twelve authors have published more than 20 articles. According to the total number of articles, Niyato (46 publications), Poor (38 publications), Han (32 publications), and Guizani (30 publications) are the most contributing authors. Yang and Liu stand out in the number of citations. Especially the article they published in 2019 received 3991 citations. These authors' TC/TP ratios are 210.74 and 191.88, respectively. Another critical performance measure is the h index. Niyato and Poor are the authors with the highest index. When the number of publications and the h index are analyzed together, the h index performance of Xiong, Zheng, and Bennis is also quite impressive.

Institutions' contributions to federated learning publications are given in Figure 7. Especially universities in China have significantly contributed to the FL area. Chinese Academy of Sciences ranks first in this field with 115 publications, followed by Beijing University of Posts Telecommunications with 94 publications. After China, the most productive organizations were found to be from Singapore. Nanyang Technological University has 91 articles in the field of FL. It was determined that the number of articles from the institutions in the top 10 is over 60.

**Table 1.** Most influential articles

| Rank | Authors | Year | Title | Journal | Total Citation | TC/ Year |
|---|---|---|---|---|---|---|
| 1 | Yang Q; Liu Y; Chen TJ; Tong YX | 2019 | Federated Machine Learning: Concept and Applications | ACM Transactions on Intelligent Systems and Technology | 3991 | 798.2 |
| 2 | Li T; Sahu AK; Talwalkar A; Smith V | 2020 | Federated Learning: Challenges, Methods, and Future Directions | IEEE Signal Processing Magazine | 1233 | 308.25 |
| 3 | Wang SQ; Tuor T; Salonidis T;Leung KK; Makaya C; He T; Chan K | 2019 | Adaptive Federated Learning in Resource Constrained Edge Computing Systems | IEEE Journal on Selected Areas in Communications | 815 | 163 |
| 4 | Lim WYB; Luong NC; Hoang DT; Jiao YT; Liang YC; Yang Q; Niyato D; Miao CY | 2020 | Federated Learning in Mobile Edge Networks: A Comprehensive Survey | IEEE Communications Surveys & Tutorials | 686 | 171.5 |
| 5 | Kairouz P; McMahan HB; Avent B; Bellet A; Bennis M; Bhagoji AN; ...; Than S. | 2021 | Advances and Open Problems in Federated Learning | Foundations and Trends in Machine Learning | 583 | 194.33 |
| 6 | Sattler F; Wiedemann S; Muller KR; Samek W | 2020 | Robust and Communication-Efficient Federated Learning From Non-i.i. Data | IEEE Transactions on Neural Networks and Learning Systems | 512 | 128 |
| 7 | Wang XF; Han YW; Wang CY; Zhao QY; Chen X; Chen M | 2019 | In-Edge AI: Intelligentizing Mobile Edge Computing, Caching, and Communication by Federated Learning | IEEE Network | 465 | 93 |
| 8 | Chen MZ; Yang ZH; Saad W; Yin CC; Poor HV; Cui SG | 2021 | A Joint Learning and Communications Framework for Federated Learning Over Wireless Networks | IEEE Transactions on Wireless Communications | 447 | 149 |
| 9 | Reke N; Hancox J; Li W; Milletari F; Roth HR; Albarqouni S; ... ;Cardoso MJ. | 2020 | The Future of Digital Health with Federated Learning | NPJ Digital Medicine | 425 | 106.25 |
| 10 | Lu YL; Huang XH; Dai YY; Maharjan S; Zhang Y | 2020 | Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT | IEEE Transactions on Industrial Informatics | 419 | 104.75 |

**Table 2.** Most productive authors

| Rank | Author | ORCID/DBLP | TP | TC | TC/TP | h_index |
|---|---|---|---|---|---|---|
| 1 | Niyato | | 46 | 2624 | 57.04 | 20 |
| 2 | Poor | | 38 | 2339 | 61.55 | 19 |
| 3 | Han | | 32 | 803 | 25.09 | 11 |
| 4 | Guizani | | 30 | 649 | 21.63 | 12 |
| 5 | Yang | | 27 | 5690 | 210.74 | 11 |
| 6 | Liu | | 25 | 4797 | 191.88 | 8 |
| 7 | Xiong | | 25 | 1261 | 50.44 | 14 |
| 8 | Hong | | 23 | 1057 | 45.96 | 11 |
| 9 | Li | | 23 | 946 | 41.13 | 9 |
| 10 | Bennis | | 22 | 1469 | 66.77 | 12 |
| 11 | Zheng | | 21 | 1385 | 65.95 | 13 |
| 12 | Quek | | 21 | 1030 | 49.05 | 10 |

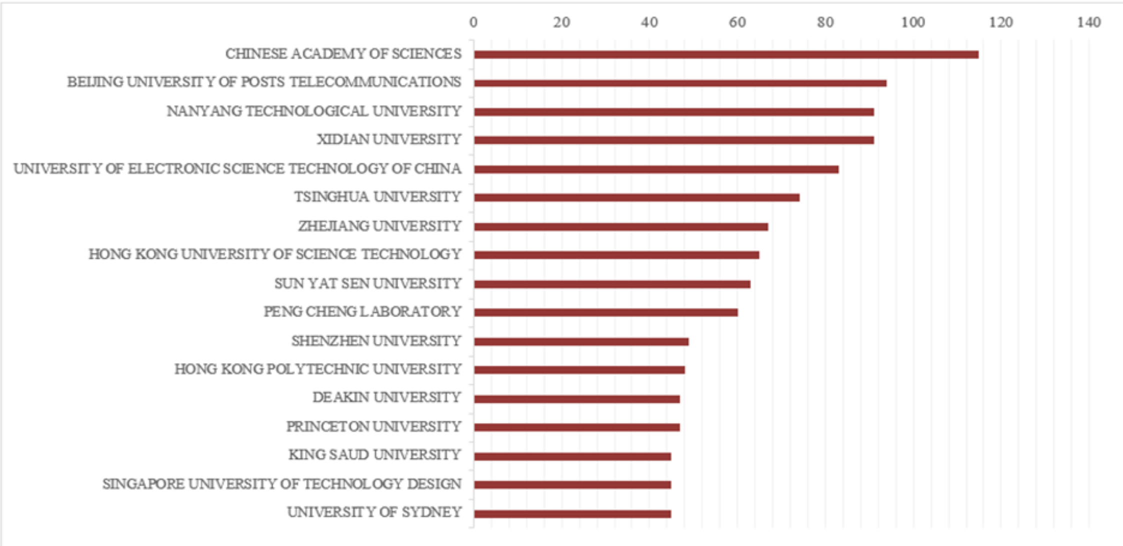



**Figure 7.** Most productive institutions

All scientific data must use the SI system (Système International d'Unités). There should be no space between the % sign and the number. The percent sign (%) is used after the number, e.g., 18%. A decimal point must be used in decimal numbers, e.g., 2.5 instead of 2,5.

## 2.2.2. Country / Region Analysis

Within the framework of bibliometric analysis, the 10 most productive countries were analyzed, and the total number of publications by year is shown in Figure 8. In terms of total articles, China (1715 articles) and the USA (477 articles) were reported to be the most significant contributors to federated learning, followed by Australia (266 articles), Canada (252 articles), England (251 articles), and South Korea (231 articles). India (176 articles), Singapore (147 articles), Saudi Arabia (126 articles), and Japan (1712 articles) countries were found to be less efficient. There has been a rapid increase in the number of publications in China over the years. In 2023 (up to August), they published 597 publications. China contributed 56% of the articles. The USA significantly contributed to federated learning with 233 publications in 2022. The publications in the USA represent 16% of the total number of articles.

The USA conducted its first article on federated learning in 2018. It is the only study in this field this year. Most of the countries in the top 10 appeared to have started work on articles on federated learning in 2019. Studies in India, which is among the top 10 most productive countries, were found to have started in 2021.

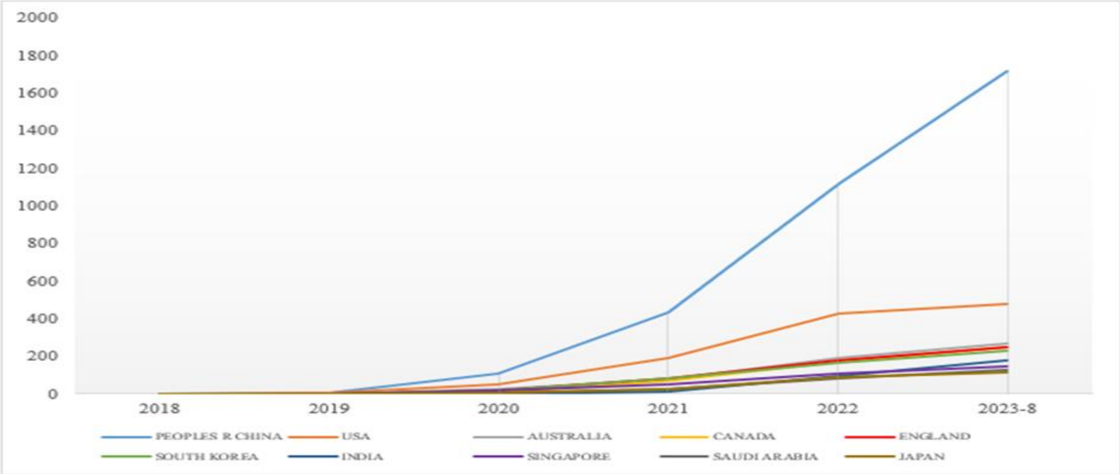



**Figure 8.** Most productive countries

## 2.2.3. Document Sources and Funding Agencies Analysis

The list of the top 10 publication titles is shown in Figure 9. It was determined that the most published journal of federated learning studies was the IEEE Internet of Things Journal, with 232 articles. IEEE Access follows this journal with 163 articles and IEEE Transactions on Industrial Informatics with 110. The number of publications in the other 10 journals in the top 10 is below 70. These 10 journals, which contribute the most to the field of FL, have 31% of the total articles.

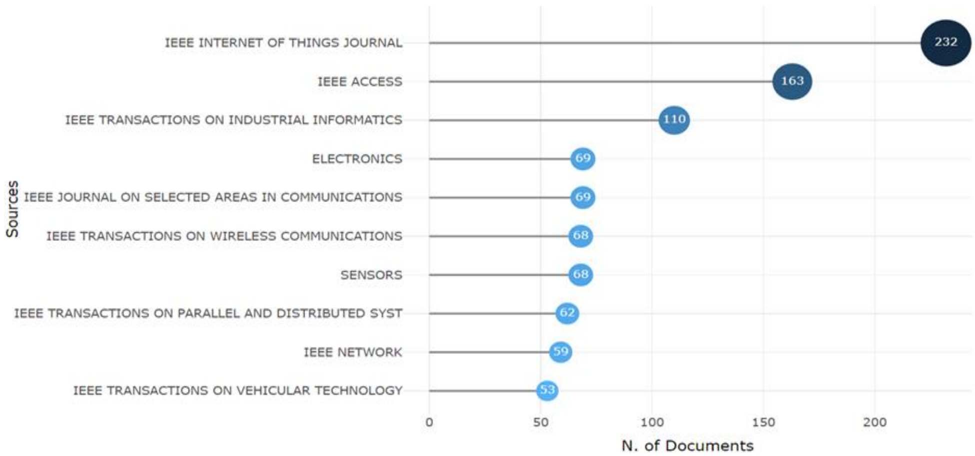



**Figure 9.** List of the top 10 contributing journals

The list of the top contributing financial agencies is shown in Figure 10. In particular, the National Natural Science Foundation of China (NSFC) ranks at the top in this field, contributing to 1112 articles. NSFC contributed more to the number of publications than the other funding agencies in the top 8. The NSFC organization was found to have funded 36% of the articles analyzed. The National Key Research and Development Program of China and Fundamental Research Funds for the Central Universities are other influential organizations in China with financial support for 242 and 161 articles, respectively. The National Science Foundation NSF (211 publications) from the USA, the National Research Foundation of Korea (126 publications), and the Ministry of Science Ict Msit Republic of Korea (96 publications) from South Korea are other prominent organizations. The European Union Eu is another organization among the top 10 funding agencies, contributing 87 articles.

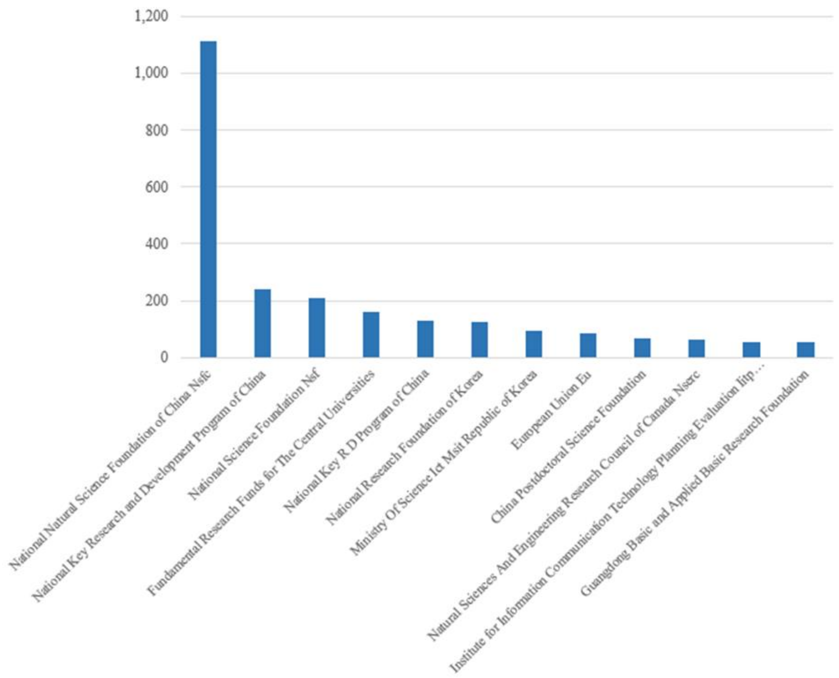



**Figure 10.** List of top contributing funding agencies

## 2.2.4. Timeline Analysis

The bibliometric analysis investigated the evolution of federated learning research content over time and produced timelines. The timeline illustrating the most commonly used terms in the titles of articles on federated learning is presented in Figure 11. In 2021-2023, it was determined that the most widely used bigrams in article titles other than federated learning were edge computing, learning framework, learning approach, machine learning, and deep learning. In 2023, the use frequency of these terms in titles was 39, 41, 27, 16, and 24. In particular, it was determined that the terms in the top 10 were used 60 or more times in article titles for three years. Upon analysis of the 8 months in 2023, it was found that the usage of the phrases reinforcement learning, intrusion detection, deep reinforcement, and mobile edge increased by 55%, 61%, 100%, and 55%, respectively, compared to 2022.

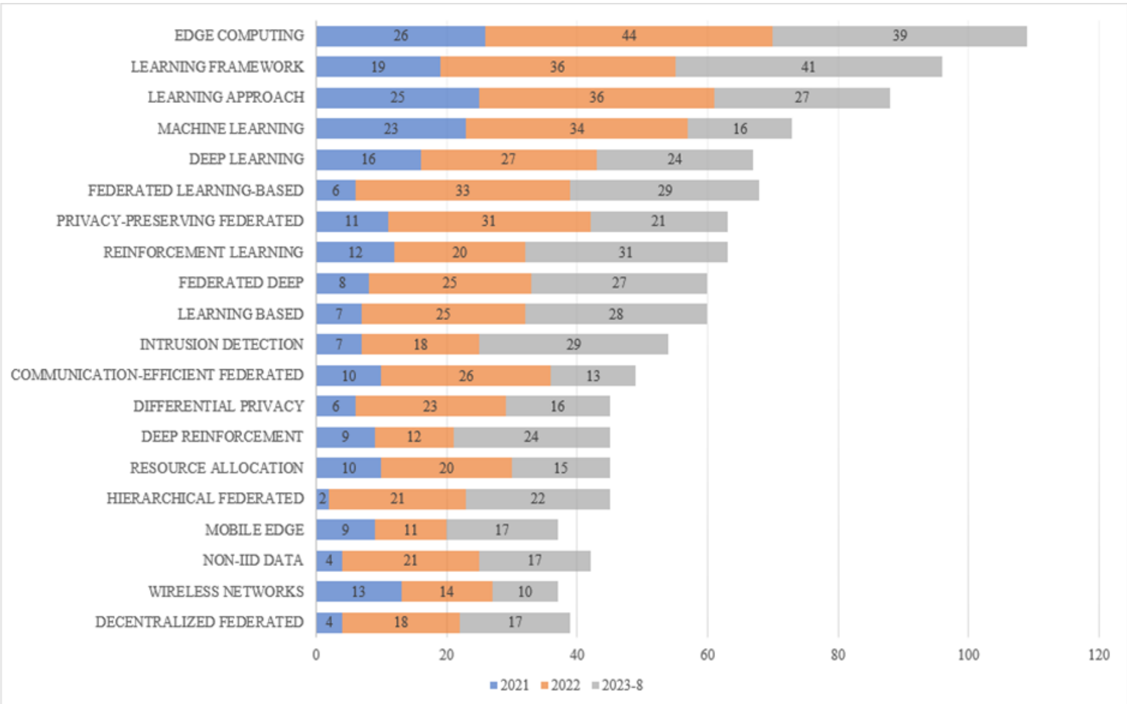
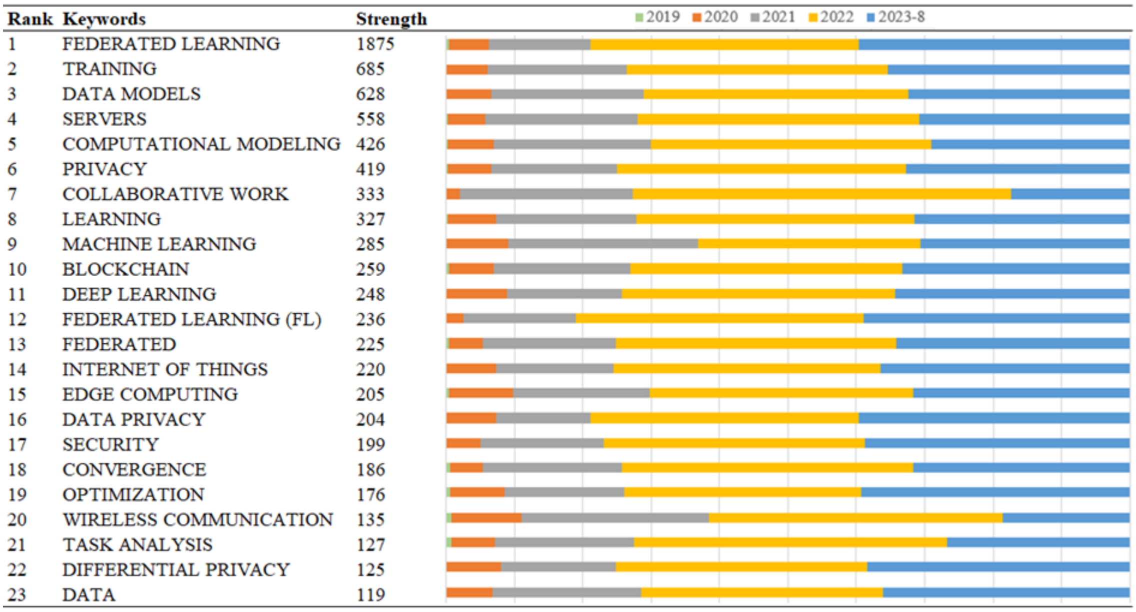



**Figure 11.** The most common bigrams in article titles




| Rank | Keywords | Strength |
|---|---|---|
| 1 | FEDERATED LEARNING | 1875 |
| 2 | TRAINING | 685 |
| 3 | DATA MODELS | 628 |
| 4 | SERVERS | 558 |
| 5 | COMPUTATIONAL MODELING | 426 |
| 6 | PRIVACY | 419 |
| 7 | COLLABORATIVE WORK | 333 |
| 8 | LEARNING | 327 |
| 9 | MACHINE LEARNING | 285 |
| 10 | BLOCKCHAIN | 259 |
| 11 | DEEP LEARNING | 248 |
| 12 | FEDERATED LEARNING (FL) | 236 |
| 13 | FEDERATED | 225 |
| 14 | INTERNET OF THINGS | 220 |
| 15 | EDGE COMPUTING | 205 |
| 16 | DATA PRIVACY | 204 |
| 17 | SECURITY | 199 |
| 18 | CONVERGENCE | 186 |
| 19 | OPTIMIZATION | 176 |
| 20 | WIRELESS COMMUNICATION | 135 |
| 21 | TASK ANALYSIS | 127 |
| 22 | DIFFERENTIAL PRIVACY | 125 |
| 23 | DATA | 119 |

**Figure 12.** The most frequently used keywords in articles

The timeline of the most frequently used keywords in articles between 2019 and 2023 is shown in Figure 12. It was determined that the most commonly used keywords in the articles were Federated Learning, Training, Data Models, and Servers. It is seen that all 23 terms in the list are used as keywords in more than 100 articles. In 2022, the usage rates of Federated Learning and Training keywords in articles were 59% and 21%, respectively. However, in 2023, there was an increase in these rates, with Federated Learning rising to 69% and Training reaching 22%. In 2023, the frequency of server, computational modeling, privacy, and collaborative work in publications declined from 18%, 14%, 14%, and 14%, respectively, to 16%, 11%, 12%, and 5%.

Academic studies must contribute towards sustainable development goals and address environmental, social, and economic challenges. The analysis of the articles within the scope of Sustainable Development Goals is shown in Figure 13. Most of the studies are within the scope of Good Health and Well-Being and Sustainable Cities and Communities. In 2022, 55% of federated learning publications were within the scope of the Good Health and Well Being sustainable target, while in 2023, this rate decreased to 37%. Federated learning applications for Industry Innovation and Infrastructure, Affordable and Clean Energy, Climate Action, Quality Education, and Responsible Consumption and Production are insufficient.

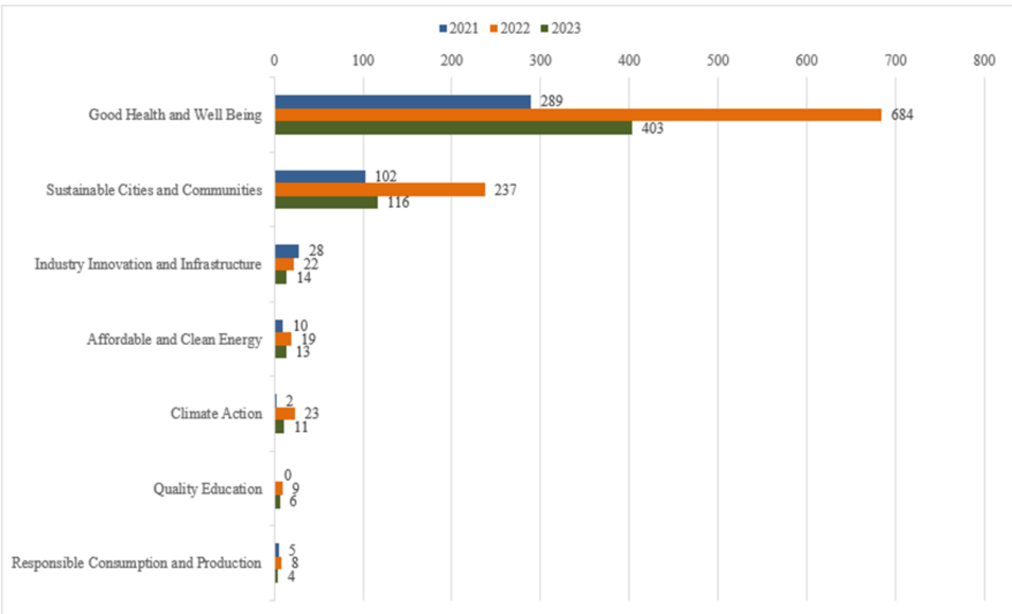



**Figure 13.** Sustainable Development Goals analysis

### 2.2.5. Social Network Analysis

The cooperation relationships of authors, organizations, and countries in federated learning article studies are given in Figures 14-16. The figures were generated using the R programming language's bibliometric library and the Shiny package framework. A link between two nodes signifies a connection between authors, countries, or organizations, where the width of the link represents the frequency of collaboration. The size of the node indicates the total link strength (TLS). Authors and institutions were categorized into six clusters, while countries were classified into two clusters differentiated by color.

In addition to examining clusters, collaborations were assessed based on betweenness and closeness values. Betweenness centrality quantifies the frequency at which a node appears on the quickest path between other nodes. In contrast, closeness centrality evaluates each node based on its proximity to all other nodes within the network.

According to the cooperation of the authors in Figure 14, Niyato's TLS was superior to the other authors. The most substantial collaboration relationship between Niyato was found to be with Xiong (18 articles), Kang (15 articles), and Wei (11 articles). Betweenness and closeness scores of Niyato were calculated to be 241.25 and 0.0145, respectively. Another author with high link strength was found to be Poor. The author frequently collaborated with Li (13 articles) and Chen (10 articles). Betweenness and closeness scores of the Poor are 97.52 and 0.013, respectively. In the analysis of author collaboration, it was observed that Han, Liu, and Bennis exhibited significant TLS values.

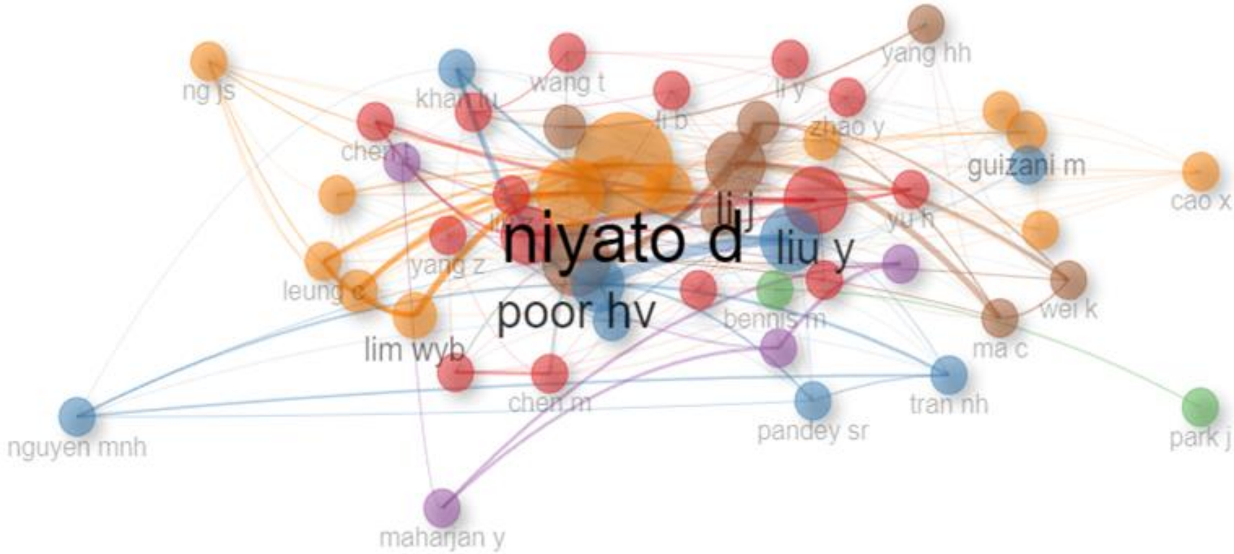



**Figure 14.** The collaboration network of authors

According to the collaborative network of institutions shown in Figure 15, the institutions with the highest betweenness scores are Nanyang Technological University, Tsinghua University, and Xidian University. The betweenness scores of these organizations are 88.22, 62.84, and 46.80, respectively. When looking at the organizations with the most collaborations, the Chinese Academy of Sciences (503 links), Nanyang Technological University (364 links), Tsinghua University (275 links), Xidian University (269 links), and Beijing University of Posts Telecommunications (202 links) stand out. The Chinese Academy of Sciences and the University of Science Technology of China Cas co-authored 39 articles, the highest collaboration.

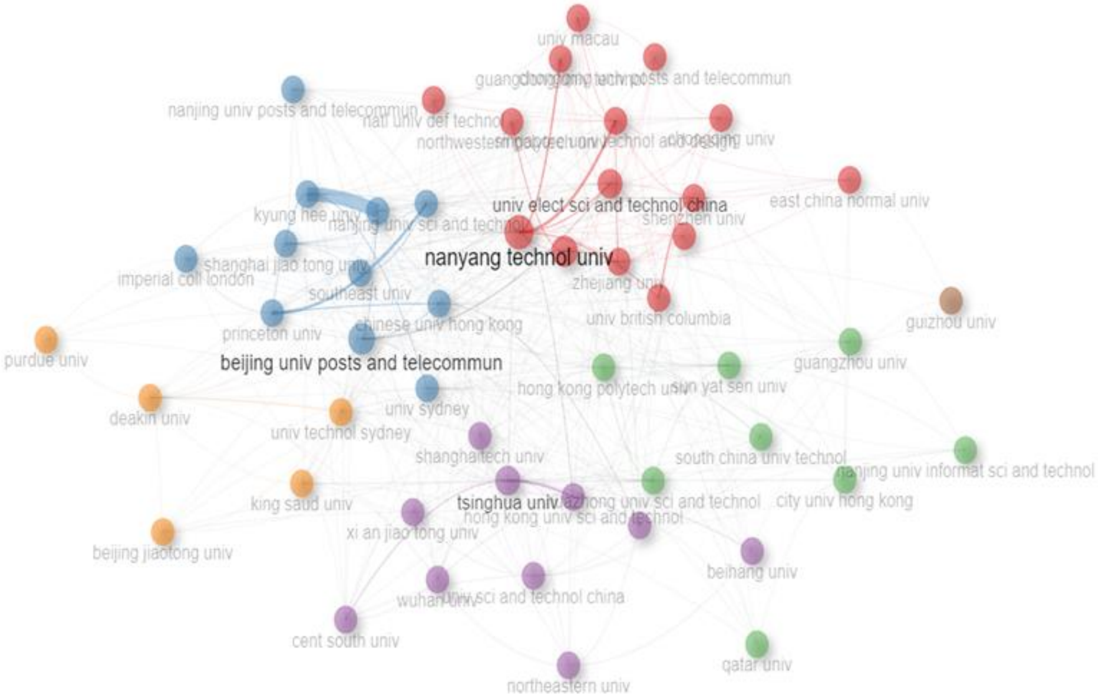



**Figure 15.** The collaboration network of institutions

The analysis reveals that in collaborations among the countries presented in Figure 16, China (1351 links), the USA (741 links), the United Kingdom (478 links), Canada (421 links), Australia (406 links), and South Korea (333 links) stand out in terms of the number of their joint publications. It was determined that China and the USA cooperated in 279 articles on federated learning. This represents 16% and 58% of the total published by China and the USA, respectively. China also has more than 100 collaborative publications with each of the following countries: the United Kingdom, Australia, Canada, and Singapore. Cooperation between the USA and Australia resulted in 56 articles, while cooperation between the USA and South Korea resulted in 55 articles.

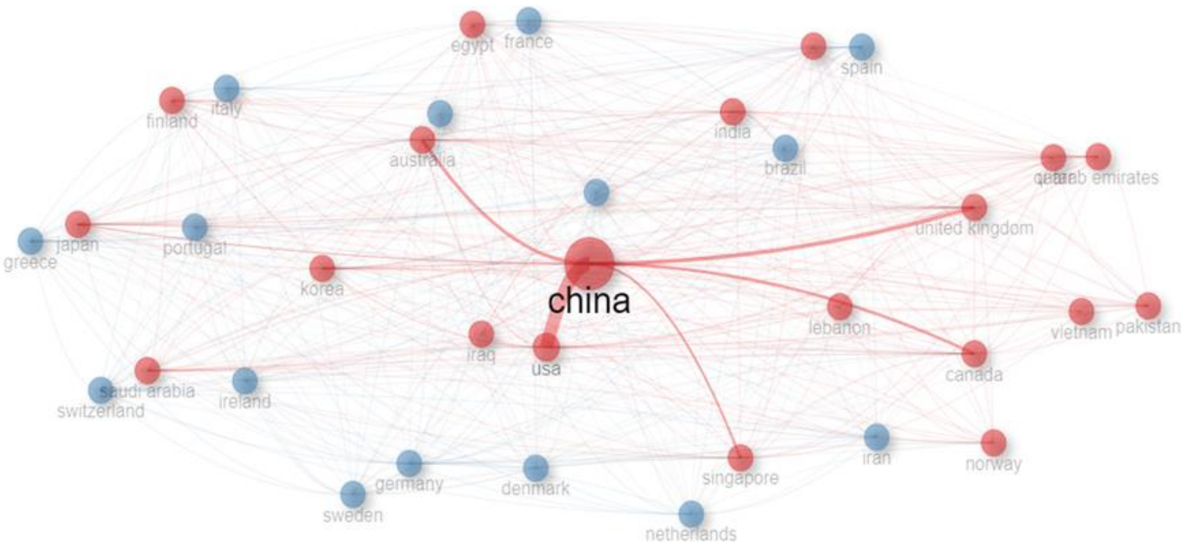



**Figure 16.** The collaboration network of countries

### 2.2.6. Keywords Detection

The bibliometric analysis also analyzed Keywords Plus data obtained by WoS from articles. Keywords Plus are words or phrases that appear frequently in the titles of an article's references but not in the article's title. The word cloud formed by the Keywords Plus obtained from the articles is shown in Figure 17. Internet, networks, and privacy are the most frequently used terms, with frequencies of 256, 189, and 169, respectively. The frequency of use of the keywords plus optimization, framework, challenges, design, and communication was over 100. Keywords such as differential privacy, wireless, big data, 5G, intrusion detection, and energy have lower frequencies.


**Figure 17.** Keywords Plus with the highest frequency

Figure 18 displays the trigram word clouds derived from the abstracts of the articles. The study revealed that the most frequently occurring trigrams within article abstracts included federated learning fl (1172 times), machine learning ml (239 times), federated learning framework (200 times), distributed machine learning (198 times), machine learning models (169 times), and artificial intelligence ai (159 times). Apart from the generally used trigrams, the most frequently used particular trigrams are identically distributed non-iid, mobile edge computing, convolutional neural network, edge computing mec, local model updates, stochastic gradient descent, and uncrewed aerial vehicles. The frequency of these trigrams was found to be more than 50.

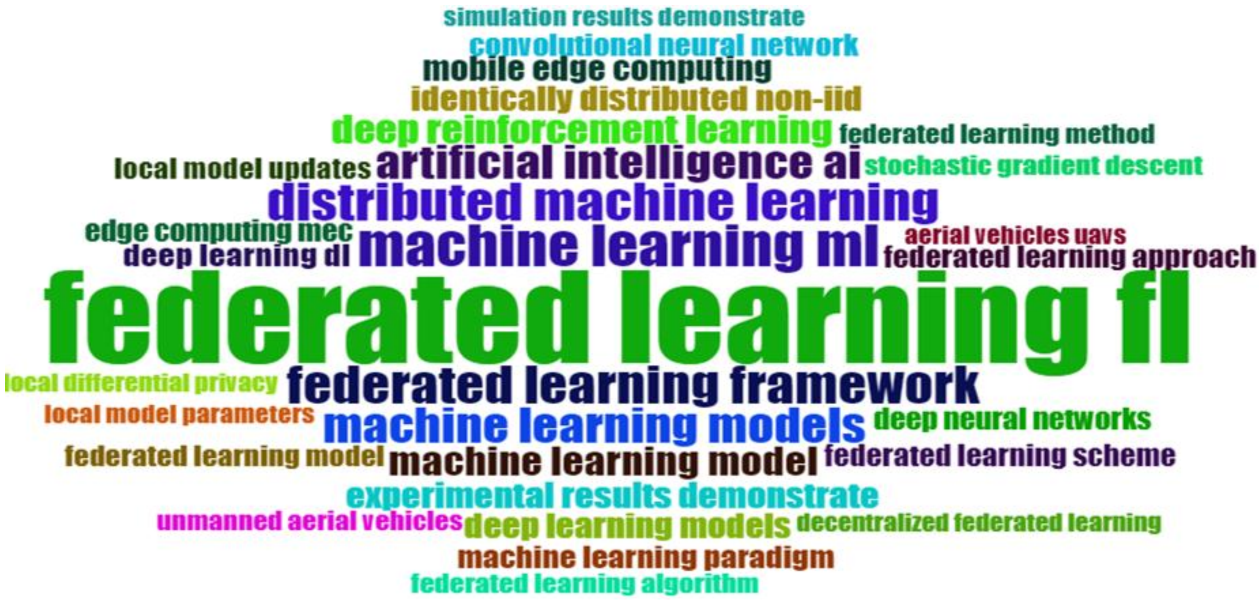



**Figure 18.** The most frequently used trigrams in article abstract

### 2.2.7. Conceptual Structure

This study section presents conceptual structure analyses of studies on federated learning. Thematic maps are intuitive plots that provide insight into the field's current and future sustainability. Themes are derived from the words in the publications. In thematic analysis, themes are derived by using words from publications. In thematic maps, these themes are characterized according to their intensity and centrality. Density is plotted on the y-axis, and centrality on the $x$-axis. Centrality gives the degree of relationship between topics; density measures the connectedness between nodes. These two characteristics measure whether specific topics are essential or not and the ability of the topic to develop and sustain itself. Themes in thematic maps can be analyzed according to the quadrant in which they are placed: (1) high density-high centrality: motor themes; (2) low density-high centrality: basic themes; (3) low density-low centrality: emerging or disappearing themes; (4) high density-low centrality: very specialized/niche themes.

Figure 19 illustrates the thematic map obtained using the keywords within the federated learning articles. The thematic map's basic themes include convergence, classification, deep, and costs. The basic themes have undergone extensive research and are challenging to anticipate their future direction. Motor themes representing advanced topics include channel estimation, signal processing, forecasting, smart grid, data privacy, optimization, and wireless communication. Distributed data, attention, security, privacy, and intelligence have been recognized as specialized/niche themes, indicating thoroughly researched but disconnected areas. Secret sharing, fault diagnosis, active learning, vertical federated learning, and poisoning attacks are low-density and centrality topics. These themes represent emerging topics or will disappear from the field of study. The federated learning theme is both a basic and motor theme.

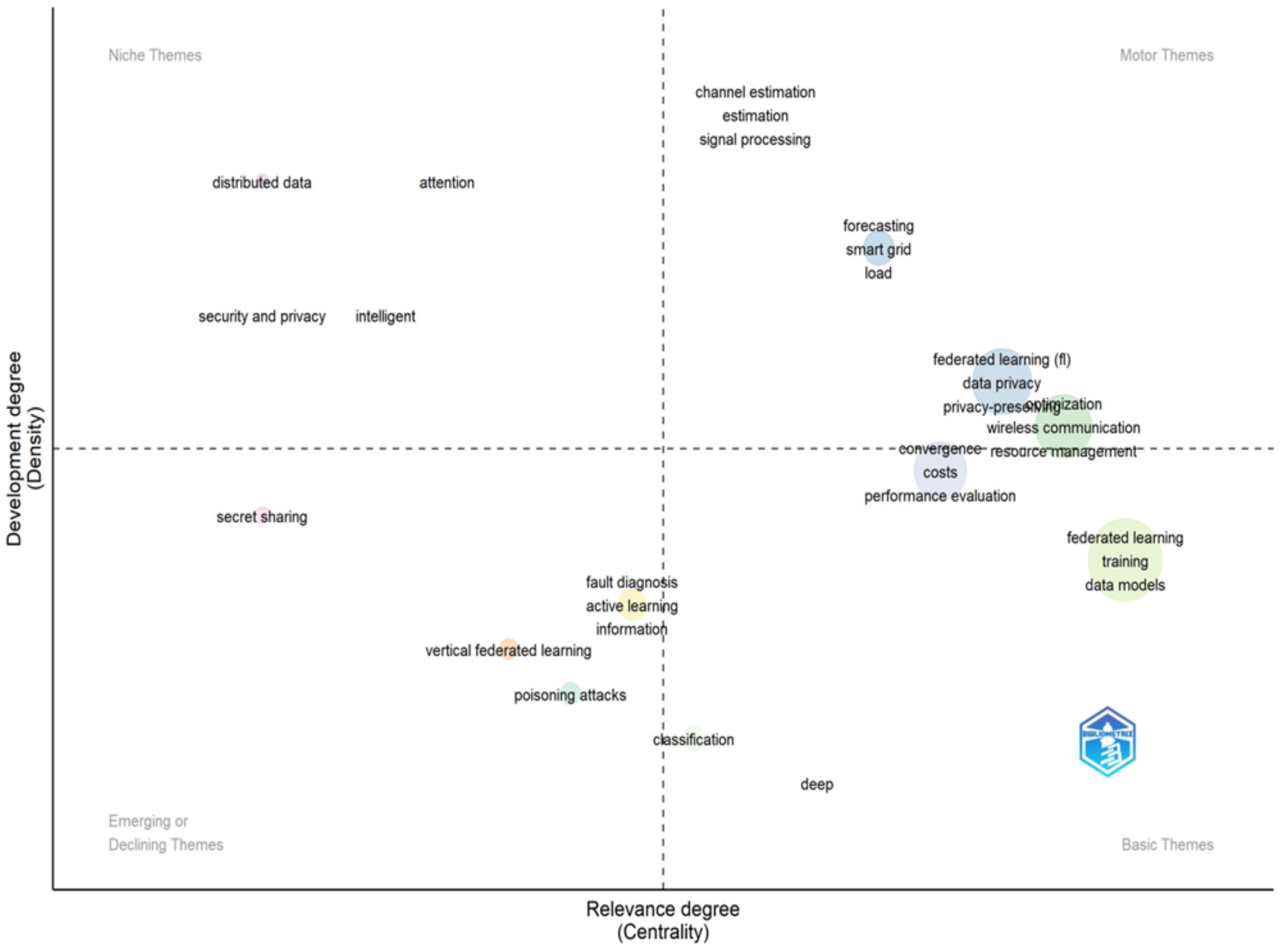



**Figure 19.** Keywords thematic map

Based on the analysis of an article titled Bigrams, the thematic map is illustrated in Figure 20. The basic themes are federated learning, deep learning, transfer learning, fault diagnosis, detection system, detection system, multiple access, and transportation systems. At the same time, generative adversarial and adversarial networks are motor themes. While gradient descent, activity recognition, image classification, and access networks are among the niche themes open to development, medical data is identified on the map as a topic that may evolve or disappear.

Factorial analysis was performed to reduce the dimensionality of the data and represent it in a low-dimensional space. Multiple correspondence analysis (MCA) is a popular approach for reducing dimensions to analyze the relationship pattern of various categorical dependent variables. The conceptual structure map generated by applying the MCA method is shown in Figure 21. The topics were clustered by determining the number of clusters as five. Cluster 1 (yellow) contains topics mostly about privacy methods, whereas Cluster 2 (blue) is about learning approaches. Cluster 3 (red) includes model-related issues like cloud computing, IoT, servers, and predictive models. Cluster 4 (green) is relevant to application areas such as resource allocation, wireless communication, energy consumption, and scheduling. A single cluster until they reach the topic in use. Height measures the distance between terms or clusters; distant terms describe different concepts.

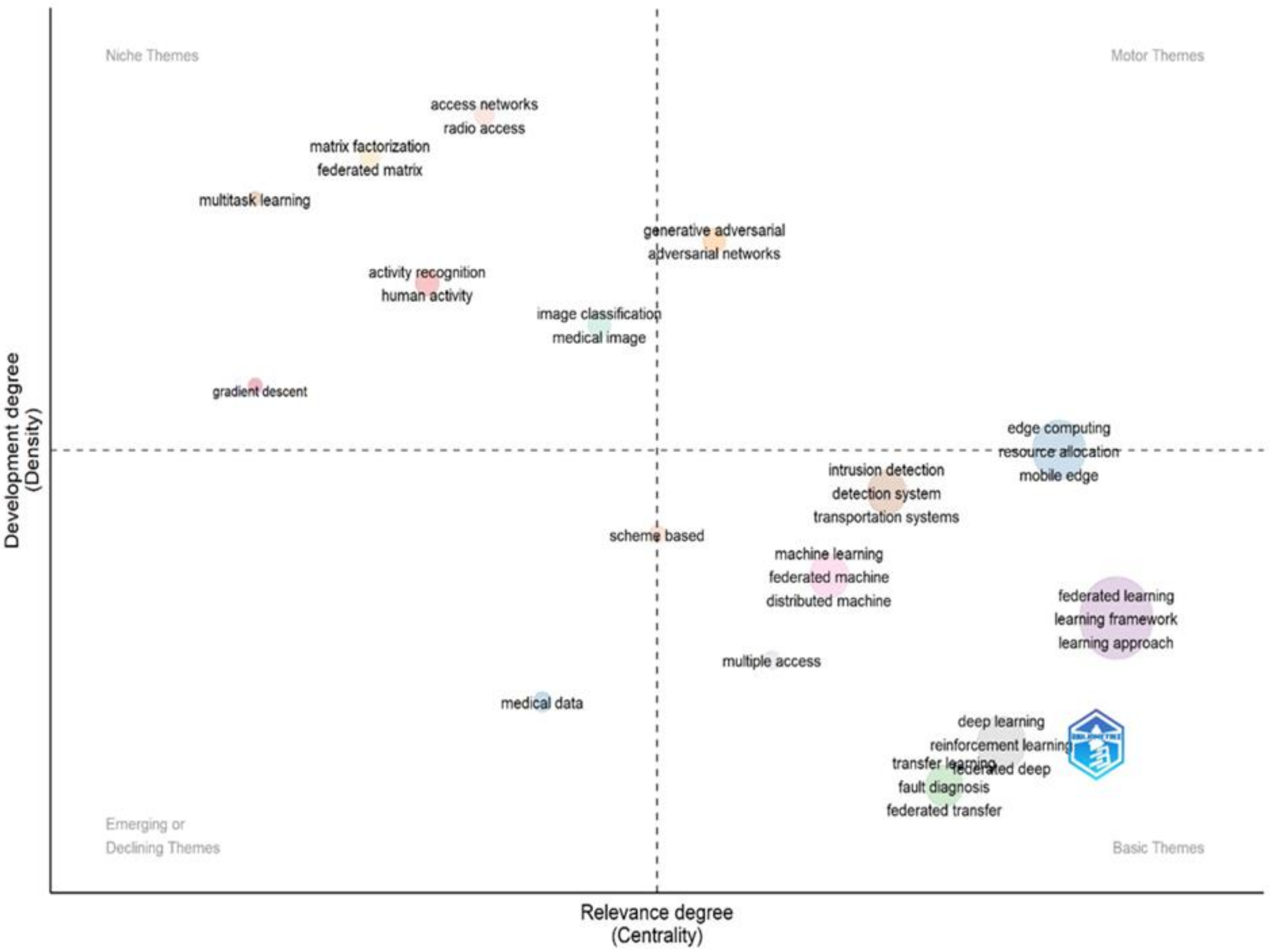



**Figure 20.** Article title bigrams thematic map

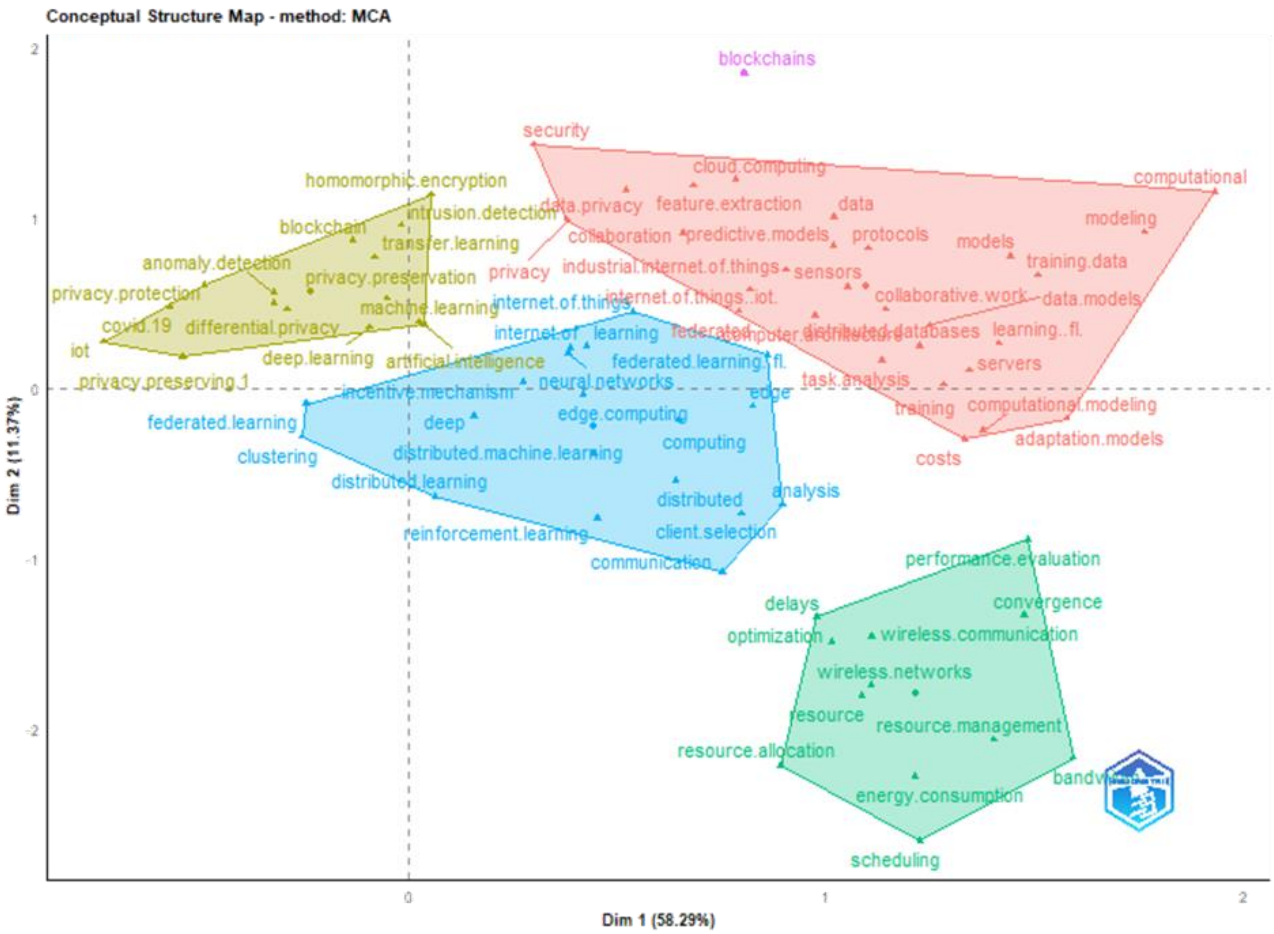


**Figure 21.** Keywords conceptual structure map

## 3. Results and Discussion

The present study analyzed articles on federated learning from the first publication until now. Publications exported from the WoS database underwent analysis, incorporating various statistical and bibliometric methods. These articles represent 54% of the total publications in different documents. Since federated learning is a new technology, the annual number of articles in this field is not very high. Still, the increase in the number of articles each year is remarkable. Our study has shown that federated learning studies are concentrated in some areas and applications, particularly in the health and energy sectors.

Analyzing publications, citations, and keywords provides insights into the trends and characteristics of articles. Between 2017 and 2023-8, six articles with 500 or more citations out of 3107 publications were retrieved from the WoS database. Niyato, Poor, Han, and Guizani are the most influential authors in the article studies of federated learning. The most commonly occurring terms as Keywords Plus were identified as Internet, Networks, and Privacy. Publications regarding a novel machine learning approach to federated learning technology are anticipated to experience exponential growth soon. In the future, it aims to carry out a study on bibliometric analysis of federated learning publications in health or energy.

The countries with the most extensive articles include China, the USA, Australia, Canada, and England. China was found to have more than 100 articles in collaboration with the following countries: the USA, United Kingdom, Australia, Canada, and Singapore. In other countries, this technology has not yet attracted much interest. Chinese Academy of Sciences (China), Beijing University of Posts Telecommunications (China), and Nanyang Technological University (Singapore) were determined as the most productive institutions.

An analysis of the topic distributions and clusters of studies in the federated learning literature shows that all studies can be grouped into four or five clusters. In particular, when interpreting the results of the conceptual structure map and the topic dendrogram, it can be seen that the study topics related to establishing data security and privacy methods, one of the main objectives of federated learning, form a cluster. Algorithms and approaches related to machine learning, deep learning, reinforcement, and distributed machine learning, which can be categorized under learning algorithms, form another cluster. As in all machine learning studies, the raw material of federated learning power is data. As a result of bibliometric analysis, a cluster of topics such as IoT, cloud computing, training data, and predictive models have been created to collect, transfer, and process data. In federated learning studies, hardware and software resources and Internet and energy-related resources are extremely important. At this point, research on planning, managing, and evaluating software, hardware, the internet, and energy resources also attracts attention as studies that form another cluster.

The initial studies on federated learning surfaced in 2017, and since then, the number of studies in this field has rapidly increased. Figures 19 and 20 provide thematic maps, and Figure 12 shows the most frequently used keywords in articles, offering insight into the course of these studies. Upon analyzing the usage rates of keywords between 2017 and 2022, it is evident that there has been an increase in the usage of specific keywords in the last two years. After analyzing the positions and potentials of keywords in thematic maps created based on keywords and headings, five main concepts emerge: security and privacy, distributed data, data privacy, data models, and optimization. Although the number of new studies and recent growth potential is increasing, specific areas of interest may decrease. These areas include data poisoning attacks, fault diagnosis, active learning information, and collaborative work on data.

Data reliability is crucial for success in machine learning applications. Traditional approaches are vulnerable to data poisoning, negatively affecting success by compromising data integrity and causing bias in training results. FL, originally conceived to establish a private and secure environment for distributed machine learning model training, currently grapples with notable privacy and security challenges. Aggregating clients' updated models on the global server is susceptible to security lapses, facilitating potential attacks such as poison attacks. Unauthorized access to updates could allow a malicious user to reverse weights, compromising client data.

Consequently, there is a burgeoning interest in researching secure and privacy-preserving FL involving homomorphic encryption, differentiable privacy, and blockchain integrated with client selection models. Preventing malicious attacks from inserting faulty data is necessary to maintain data integrity. Poisoning attack modes, such as model distortion and feedback weaponization, are a topic of study with potential implications for federated learning.

Optimization is one of the most common areas for future work in FL. The optimal selection of customers is crucial in FL because model updates enable global model convergence. Existing customer selection models prioritize how to select customers rather than determining the required number. Currently, a fixed number of customers without standardized criteria is widely used. Although increasing the number of clients may improve convergence, practical limitations such as resource constraints prevent this approach. Therefore, research is needed to determine the optimal number of clients. Additionally, further research is required to develop mechanisms that can dynamically determine client selection in each FL round.

The development of technology and the decrease in sensitivity towards data privacy have caused major problems. Federated learning technology addresses customers' concerns in this regard. Not only is learning a promising technology federated, but it is also a new artificial intelligence business model. Especially considering its advantages of privacy and security, scalability, low power requirements, improved accuracy of results, reduced data training costs, and data minimization, this approach is likely to become increasingly common in machine learning research. On the other hand, the current federated learning technology has many problems related to communication and security breaches. Improvements in the methodology are significant to solve these problems. Therefore, the number of publications on topics such as local training, merging models, communication of clients, communication costs, and attacks on the system will increase.

Once federated learning methodologies are developed, collaborative work on data analysis using FL technology by companies in the same country or different countries will increase rapidly. In particular, companies with varying service areas can overcome the lack of customer information with this approach and achieve higher percentage accuracy rates in classification, clustering, and prediction studies. Federated learning will become a major requirement in healthcare, energy, retail, smart transport, insurance, banking, manufacturing, etc.

## 4. Conclusion

This study provides an extensive and detailed bibliometric analysis of FL publications, offering valuable insights into the evolution of research in this field. The results show that China and the USA dominate publications, while FL applications in health and energy sectors are emerging as significant focus areas. The thematic analysis also reveals strong research clusters around data privacy, machine learning approaches, and optimization techniques. These findings demonstrate FL's rapid growth and diverse applications, particularly in addressing privacy concerns.

Future studies should address the security and communication challenges associated with FL. As privacy-preserving methods such as homomorphic encryption and differential privacy gain traction, further research is needed to improve scalability and reduce communication costs. Additionally, developing dynamic client selection mechanisms and more efficient optimization techniques will be essential to enhance FL's effectiveness in real-world applications across various industries.

## Author Contributions

The first author: Conceptualization, software, formal, analysis writing—review and editing, visualization. The second author: Data curation, visualization, supervision, project administration. The third author: Methodology, validation, investigation, writing—original draft preparation, supervision. The fourth author: Resources data curation, project administration. All authors read and approved the final version of the paper.

## Conflicts of Interest

All the authors declare no conflict of interest.

## Ethical Review and Approval

No approval from the Board of Ethics is required.

## References

[1] L. Li, Y. Fan, M. Tse, K. Y. Lin, *A review of applications in federated learning*, Computers and Industrial Engineering 149 (2020) 106854 15 pages.

[2] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, Y. Gao, *A survey on federated learning*, Knowledge-Based Systems 216 (2021) 106775 11 pages.

[3] C. Yang, Q. Wang, M. Xu, Z. Chen, K. Bian, Y. Liu, X. Liu, *Characterizing impacts of heterogeneity in federated learning upon large-scale smartphone data*, in: J. Leskovec, M. Grobelnik, M. Najork, J. Tang, L. Zia (Eds.), Proceedings of the Web Conference 2021, Ljubljana, 2021, pp. 935-946.

[4] A. Imteaj, U. Thakker, S. Wang, J. Li, M. H. Amini, *A survey on federated learning for resource-constrained IoT devices*, IEEE Internet of Things Journal 9 (1) (2021) 1-24.

[5] S. Savazzi, M. Nicoli, V. Rampa, *Federated learning with cooperating devices: A consensus approach for massive IoT networks*, IEEE Internet of Things Journal 7 (5) (2020) 4641-4654.

[6] J. Pang, Y. Huang, Z. Xie, Q. Han, Z. Cai, *Realizing the heterogeneity: A self-organized federated*

*learning framework for IoT*, IEEE Internet of Things Journal 8 (5) (2020) 3088-3098.

[7] A. Qayyum, K. Ahmad, M. A. Ahsan, A. Al-Fuqaha, J. Qadir, *Collaborative federated learning for healthcare: Multi-modal COVID-19 diagnosis at the edge*, IEEE Open Journal of the Computer Society 3 (2022) 172–184.

[8] L. Sun, J. Wu, *A scalable and transferable federated learning system for classifying healthcare sensor data*, IEEE Journal of Biomedical and Health Informatics 27 (2) (2022) 866–877.

[9] A. Jiménez-Sánchez, M. Tardy, M. A. G. Ballester, D. Mateus, G. Piella, *Memory-aware curriculum federated learning for breast cancer classification*, Computer Methods and Programs in Biomedicine 229 (2023) 107318 10 pages.

[10] J. Bian, J. Huang, S. Ji, Y. Liao, X. Li, Q. Wang, J. Zhou, D. Dou, Y. Wang, H. Xiong, *Feynman: Federated learning-based advertising for ecosystems-oriented mobile apps recommendation*, IEEE Transactions on Services Computing 16 (5) (2023) 3361–3372.

[11] S. R. Pokhrel, J. A. Choi, *Decentralized federated learning approach for connected autonomous vehicles*, in: S. Bahk, Y. Kim, S. C. Bang, S. L. Kim, Y. J. Choi (Eds.), IEEE Wireless Communications and Networking Conference Workshops, Seoul, 2020, pp. 1–6

[12] T. Zeng, O. Semiari, M. Chen, W. Saad, M. Bennis, *Federated learning on the road autonomous controller design for connected and autonomous vehicles*, IEEE Transactions on Wireless Communications 21 (12) (2022) 10407–10423.

[13] Y. M. Saptura, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, S. Srikanteswara, *Energy demand prediction with federated learning for electric vehicle networks*, in: V. Cerf, D. Zuckerman, T. Mooren, J. Hong (Eds.), IEEE Global Communications Conference, Waikoloa, 2019, pp. 1–6

[14] M. Savi, F. Olivadese, *Short-term energy consumption forecasting at the edge: A federated learning approach*, IEEE Access 9 (2021) 95949–95969.

[15] Y. Wang, I. L. Bennani, X. Liu, M. Sun, Y. Zhou, *Electricity consumer characteristics identification: A federated learning approach*, IEEE Transactions on Smart Grid 12 (4) (2021) 3637–3647.

[16] Q. Yang, Y. Liu, T. Chen, Y. Tong, *Federated machine learning: Concept and applications*, ACM Transactions on Intelligent Systems and Technology 10 (2) (2019) 1–19.

[17] H. Gupta, D. Patel, A. Makade, K. Gupta, O. P. Vyas, A. Puliafito, *Risk prediction in the life insurance industry using federated learning approach*, IEEE 21st Mediterranean Electrotechnical Conference, Palermo, 2022, pp. 948–953.

[18] D. Jatain, V. Singh, N. Dahiya, *A contemplative perspective on federated machine learning: Taxonomy, threats and vulnerability assessment and challenges*, Journal of King Saud University-Computer and Information Sciences 34 (9) (2022) 6681–6698.

[19] J. Li, T. Cui, K. Yang, R. Yuan, L. He, M. Li, *Demand forecasting of e-commerce enterprises based on horizontal federated learning from the perspective of sustainable development*, Sustainability 13 (23) (2021) 13050 29 pages.

[20] M. V. Shenoy, *HFedDI: A novel privacy preserving horizontal federated learning-based scheme for IoT device identification*, Journal of Network and Computer Applications 214 (2023) 103616 12 pages.

[21] H. Wang, F. Xie, Q. Duan, J. Li, *Federated learning for supply chain demand forecasting*, Mathematical Problems in Engineering 2022 (1) (2022) Article ID 4109070 8 pages.

[22] Z. Mao, H. Li, Z. Huang, Y. Tian, P. Zhao, Y. Li, *Full data-processing power load forecasting based on vertical federated learning*, Journal of Electrical and Computer Engineering 2023 (1) (2023) Article ID 9914169 9 pages.

[23] I. Kevin, K. Wang, X. Zhou, W. Liang, Z. Yan, J. She, *Federated transfer learning based cross-domain prediction for smart manufacturing*, IEEE Transactions on Industrial Informatics 18 (6) (2021) 4088–4096.

[24] Y. Chen, X. Qin, J. Wang, C. Yu, W. Gao, *Fedhealth: A federated transfer learning framework for wearable healthcare*, IEEE Intelligent Systems 35 (4) (2020) 83–93.

[25] D. Ye, R. Yu, M. Pan, Z. Han, *Federated learning in vehicular edge computing: A selective model aggregation approach*, IEEE Access 8 (2020) 23920–23935.

[26] X. Zhou, W. Liang, J. She, Z. Yan, I. Kevin, K. Wang, *Two-layer federated learning with heterogeneous model aggregation for 6g supported internet of vehicles*, IEEE Transactions on Vehicular Technology 70 (6) (2021) 5308–5317.

[27] W. Ou, J. Zeng, Z. Guo, W. Yan, D. Liu, S. Fuentes, *A homomorphic-encryption-based vertical federated learning scheme for risk management*, Computer Science and Information Systems 17 (3) (2020) 819–834.

[28] L. Zhang, J. Xu, P. Vijayakumar, P. K. Sharma, U. Ghosh, *Homomorphic encryption-based privacy-preserving federated learning in IoT-enabled healthcare system*, IEEE Transactions on Network Science and Engineering 10 (5) (2022) 2864–2880.

[29] Y. Zhao, J. Zhao, M. Yang, T. Wang, N. Wang, L. Lyu, D. Niyato, K. Y. Lam, *Local differential privacy-based federated learning for the Internet of things*, IEEE Internet of Things Journal 8 (11) (2021) 8836–8853.

[30] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, H. R. Tizhoosh, *Federated learning and differential privacy for medical image analysis*, Scientific Reports 12 (1) (2022) Article Number 1953 10 pages.

[31] D. Jatain, V. Singh, N. Dahiya, *A contemplative perspective on federated machine learning: Taxonomy, threats and vulnerability assessment and challenges*, Journal of King Saud University-Computer and Information Sciences 34 (9) (2022) 6681–6698.

[32] B. Jia, X. Zhang, J. Liu, Y. Zhang, K. Huang, Y. Liang, *Blockchain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT*, IEEE Transactions on Industrial Informatics 18 (6) (2021) 4049–4058.

[33] Z. Chen, P. Tian, W. Liao, W. Yu, *Zero-knowledge clustering based adversarial mitigation in heterogeneous federated learning*, IEEE Transactions on Network Science and Engineering 8 (2) (2020) 1070–1083.

[34] V. C. Gogineni, S. Werner, Y. F. Huang, A. Kuh, *Communication-efficient online federated learning framework for nonlinear regression*, in: H. Li, S. Furui, T. Chen (Eds.), IEEE International Conference on Acoustics, Speech and Signal Processing, Singapore, 2022, pp. 5228–5232.

[35] Q. Wei, Q. Li, Z. Zhou, Z. Q. Ge, Y. Zhang, *Privacy-preserving two-parties logistic regression on vertically partitioned data using asynchronous gradient sharing*, Peer-to-Peer Networking and Applications 14 (2021) 1379–1387.

[36] J. Cai, X. Liu, Z. Yu, K. Guo, J. Li, *Efficient vertical federated learning method for ridge regression of large-scale samples*, IEEE Transactions on Emerging Topics in Computing 11 (2) (2022) 511–526.

[37] F. Yamamoto, S. Ozawa, L. Wang, *eFL-Boost: Efficient federated learning for gradient boosting decision trees*, IEEE Access 10 (2022) 43954–43963.

[38] X. Chen, S. Zhou, B. Guan, K. Yang, H. Fao, H. Wang, Y. Wang, *Fed-eini: An efficient and interpretable inference framework for decision tree ensembles in vertical federated learning*, in: U. Fayyad, X. Zhu (Eds.), IEEE International Conference on Big Data, Virtual Event, 2021, pp. 1242–1248

[39] T. Markovic, M. Leon, D. Buffoni, S. Punnekkat, *Random forest based on federated learning for intrusion detection*, in: I. Maglogiannis, J. Macintyre (Eds.), International Conference on Artificial Intelligence Applications and Innovations, Springer, Cham, 2022, pp. 132–144.

[40] T. U. Islam, R. Ghasemi, N. Mohammed, *Privacy-preserving federated learning model for healthcare data*, in: S. Vuong (Ed.), IEEE 12th Annual Computing and Communication Workshop and Conference, Las Vegas, 2022, pp. 0281–0287.

[41] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, W. Shi, *Federated learning of predictive models from federated electronic health records*, International Journal of Medical Informatics 112 (2018) 59–67.

[42] A. Vaid, S. K Jaladanki, J. Xu, S. Teng, A. Kumar, S. Lee, S. Somani, I. Paranjpe, J. K De Freitas, T. Wanyan, K. W. Johnson, M. Bıçak, E. Klang, Y. J. Kwon, A. Costa, S. Zhao, R. Miotto, A. W. Charney, E. Böttinger, Z. A. Fayad, G. N. Nadkarni, F. Wang, B. S. Glicksberg, *Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: Machine learning approach*, JMIR Medical Informatics 9 (1) (2021) e24207 11 pages.

[43] P. Guo, P. Wang, J. Zhou, S. Jiang, V. M. Patel, *Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning*, in: M. S. Brown, R. Sukthankar, T. Tan, L. Zelnik (Eds.), Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, 2021 pp. 2423–2432.

[44] I. Khan, A. Guerrieri, G. Spezzano, A. Vinci, *Occupancy prediction in buildings: An approach leveraging LSTM and Federated Learning*, IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, Calabria, 2022 pp. 1–7.

[45] G. Zhang, S. Zhu, X. Bai, *Federated learning-based multi-energy load forecasting method using CNN-Attention-LSTM model*, Sustainability 14 (19) (2022) 12843 14 pages.

[46] S. Agrawal, S. Sarkar, M. Alazab, P. K. R. Maddikunta, T. R. Gadekallu, Q. V. Pham, *Genetic CFL: Hyperparameter optimization in clustered federated learning*, Computational Intelligence and Neuroscience 2021 (1) (2021) Article ID 7156420 10 pages.

[47] H. N. C. Neto, I. Dusparic, D. M. F. Mattos, N. C. Fernandes, *FedSA: Accelerating intrusion detection in collaborative environments with federated simulated annealing*, in: A. Clemm, G. Maier (Eds.), IEEE 8th International Conference on Network Softwarization, Milan, 2022, pp. 420–428.

[48] S. Park, Y. Suh, J. Lee, *FedPSO: Federated learning using particle swarm optimization to reduce communication costs*, Sensors 21 (2) (2021) 600 13 pages.

[49] F. Cui, Z. A. Al-Sudani, G. S. Hassan, H. A. Afan, S. J. Ahammed, Z. M. Yaseen, *Boosted artificial intelligence model using improved alpha-guided grey wolf optimizer for groundwater level prediction: Comparative study and insight for federated learning technology*, Journal of Hydrology 606 (2022) 127384 16 pages.

[50] X. Zhu, D. Wang, W. Pedrycz, Z. Li, *Horizontal federated learning of Takagi–Sugeno fuzzy rule-based models*, IEEE Transactions on Fuzzy Systems 30 (9) (2021) 3537–3547.

[51] N. Cha, Z. Du, C. Wu, T. Yoshinaga, L. Zhong, J. Ma, F. Liu, Y. Ji, *Fuzzy logic-based client selection for federated learning in vehicular networks*, IEEE Open Journal of the Computer Society 3 (2022) 39–50.

[52] E. Yoo, H. Ko, S. Pack, *Fuzzy clustered federated learning algorithm for solar power generation forecasting*, IEEE Transactions on Emerging Topics in Computing 10 (4) (2022) 2092–2098.

[53] F. A. KhoKhar, J. H. Shah, M. A. Khan, M. Sharif, U. Tariq, S. Kadry, *A review on federated learning*

*towards image processing*, Computers and Electrical Engineering 99 (2022) 107818 19 pages.

[54] E. M. Campos, P. F. Saura, A. González-Vidal, J. L. Hernández-Ramos, J. B. Bernabé, G. Baldini, A. Skarmeta, *Evaluating federated learning for intrusion detection in internet of things: Review and challenges*, Computer Networks 203 (2022) 108661 16 pages.

[55] L. G. F. da Silva, D. F. H. Sadok, P. T. Endo, *Resource optimizing federated learning for use with IoT: A systematic review*, Journal of Parallel and Distributed Computing 175 (2023) 92–108.

[56] H. Li, C. Li, J. Wang, A. Yang, Z. Ma, Z. Zhang, D. Hua, *Review on security of federated learning and its application in healthcare*, Future Generation Computer Systems 7144 (2023) 271–290.

[57] F. Liu, M. Li, X. Liu, T. Xue, J. Ren, C. Zhang, *A review of federated meta-learning and its application in cyberspace security*, Electronics 12 (15) (2023) 3295 35 pages.

[58] Z. Teimoori, A. Yassine, *A review on intelligent energy management systems for future electric vehicle transportation*, Sustainability 14 (21) (2022) 14100 23 pages.

[59] Z. Yang, M. Chen, K. K. Wong, H. V. Poor, S. Cui, *Federated learning for 6G: Applications, challenges, and opportunities*, Engineering 8 (2022) 33–41.

[60] A. R. Javed, M. A. Hassan, F. Shahzad, W. Ahmed, S. Singh, T. Baker, T. R. Gadekallu, *Integration of blockchain technology and federated learning in vehicular (IoT) networks: A comprehensive survey*, Sensors 22 (12) (2022) 4394 24 pages.

[61] H. S. Sikandar, H. Waheed, S. Tahir, S. U. R. Malik, W. Rafique, *A detailed survey on federated learning attacks and defenses*, Electronics 12 (2) (2023) 260 18 pages.

[62] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, W. M. Lim, *How to conduct a bibliometric analysis: An overview and guidelines*, Journal of Business Research 133 (2021) 285–296.

[63] M. Ulu, Y. S. Türkan, *Bibliometric analysis of traffic accident prediction studies from 2003 to 2023: Trends, patterns, and future directions*, Promet-Traffic and Transportation 36 (5) (2024) 1–21.

[64] Y. Li, Z. Xu, X. Wang, X. Wang, *A bibliometric analysis on deep learning during 2007–2019*, International Journal of Machine Learning and Cybernetics 11 (2020) 2807–2826.

[65] J. M. Merigó, J. B. Yang, *A bibliometric analysis of operations research and management science*, Omega 73 (2017) 37–48.

[66] A. E. Ezugwu, A. K. Shukla, R. Nath, A. A. Akinyelu, J. O. Agushaka, H. Chiroma, P. K. Muhuri, *Metaheuristics: A comprehensive overview and classification along with bibliometric analysis*, Artificial Intelligence Review 54 (2021) 4237–4316.

[67] A. K. Shukla, M. Janmaijaya, A. Abraham, P. K. Muhuri, *Engineering applications of artificial intelligence: A bibliometric analysis of 30 years* (*1988–2018*), Engineering Applications of Artificial Intelligence 85 (2019) 517–532

[68] J. Z. Zhang, P. R. Srivastava, D. Sharma, P. Eachempati, *Big data analytics and machine learning: A retrospective overview and bibliometric analysis*, Expert Systems with Applications 184 (2021) 115561 18 pages.

[69] D. Yu, Z. Xu, W. Wang, *A bibliometric analysis of Fuzzy Optimization and Decision Making (2002–2017)*, Fuzzy Optimization and Decision Making 18 (2019) 371–397.

[70] Y. Zhang, H. Chen, J. Lu, G. Zhang, *Detecting and predicting the topic change of knowledge-based systems: A topic-based bibliometric analysis from 1991 to 2016*, Knowledge-Based Systems 133 (2017) 255–268.

[71] Clarivate, (2023), https://clarivate.libguides.com/librarianresources/coverage, Accessed 15 Nov 2023.

[72] W. Yang, Y. Zhang, K. Ye, L. Li, C. Z. Xu, *Ffd: A federated learning based method for credit card fraud detection*, in: K. Chen, S. Seshadri, L. J. Zhang (Eds.), Big Data–BigData 2019: 8th International Congress, Held as Part of the Services Conference Federation, San Diego, 2019, Springer, Proceedings 8, pp. 18–32.