IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 February 2024, accepted 11 March 2024, date of publication 18 March 2024, date of current version 22 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3378126

SURVEY

Insights Into Privacy Protection Research in AI

SHASHA YU[1], FIONA CARROLL[2], (Senior Member, IEEE), AND BARRY L. BENTLEY[2]
[1]School of Professional Studies, Clark University, Worcester, MA 01610, USA
[2]Cardiff School of Technologies, Cardiff Metropolitan University, CF5 2YB Cardiff, U.K.

Corresponding author: Shasha Yu (ShaYu@clarku.edu)

ABSTRACT This paper presents a systematic bibliometric analysis of the artificial intelligence (AI) domain to explore privacy protection research as AI technologies integrate and data privacy concerns rise. Understanding evolutionary patterns and current trends in this research is crucial. Leveraging bibliometric techniques, the authors analyze 8,322 papers from the Web of Science (WoS) database, spanning 1990 to 2023. The analysis highlights *IEEE Transactions on Knowledge and Data Engineering* and *IEEE Access* journals as highly influential, the former being an early contributor and the latter emerging as a pivotal source. The study demonstrates substantial disparities in scientific productivity across countries. Specifically, the top 10 countries collectively accounted for 74.8% of the articles, with China and the USA making up nearly half of the total contribution (46.1%). In contrast, regions in Africa and South America exhibited lower scientific production. The evolution of privacy preservation research is reflected, shifting from an algorithm-oriented approach to a focus on data orientation, and subsequently, to privacy solutions centered around Cloud Computing. In recent years, there has been a shift towards embracing Federated Learning and Differential Privacy. The analysis brings to light emerging themes and identifies research gaps, notably a global disparity in research output and a lag in ethical and legal inquiry. It asserts that enhanced interdisciplinary collaboration is imperative to formulate comprehensive privacy solutions for AI. Specifically, the paper imparts invaluable insights that are pivotal for effectively addressing the evolving privacy concerns in the era of AI and big data.

INDEX TERMS AI, artificial intelligence, bibliometric analysis, privacy protection.

I. INTRODUCTION

In recent years, the field of artificial intelligence (AI) has witnessed a remarkable surge in development, with its profound impact extending across various domains, from healthcare [1] and finance [2] to transportation [3] and agriculture [4]. This exponential growth has not only revolutionized industries but has also become an integral part of our daily lives. AI, with its ability to process vast amounts of data and make informed decisions, has left an indelible mark on society, with the future potential to influence virtually every aspect of out existence [5].

However, amid the exhilarating advancements in AI technology, a pressing concern has emerged – the safeguarding of personal privacy in the AI era [6]. As AI systems rely extensively on data [7], the omnipresence of these technologies raises significant questions about the protection of individual privacy and the potential misuse of personal information [8]. The rapid proliferation of AI applications, including but not limited to personalized recommendations [9], autonomous vehicles [10], virtual assistants [11], and medical diagnostics [12], underscores the urgency of addressing these privacy concerns.

In November 2021, UNESCO's first-ever global standard on AI ethics, the *Recommendation on the Ethics of Artificial Intelligence*, was adopted by all 193 Member States [13]. This landmark decision underscored the imperative that privacy be respected, protected, and promoted throughout the life cycle of AI systems. Furthermore, it advocated for the establishment of robust data protection frameworks and governance mechanisms, supported by judicial systems, and maintained throughout the AI systems' lifecycle [14]. This initiative set a global normative framework, entrusting States with the responsibility to implement it at their level. By October 2023, over 50 national strategic and government-wide initiatives for trustworthy AI were in place [15], with numerous countries and organizations working to establish new policies and regulations to safeguard individuals' privacy. For instance,

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

in October 2023, President Biden of the United States issued an Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, urging Congress to enact bipartisan data privacy legislation. This order also directed actions to enhance privacy-preserving research and technologies and to develop guidelines for federal agencies to assess the effectiveness of privacy-preserving techniques [16]. The EU's proposed Artificial Intelligence Act (AI Act) is another example. It aims to address the potential risks AI systems pose to fundamental rights, including privacy and data protection. It introduces a risk-based classification system for AI, with specific requirements and obligations tailored to each category. This approach underscores the importance of privacy protection within AI applications [17].

It is against this backdrop that this paper endeavors to shed light on the critical issue of privacy protection in the realm of AI. To navigate the complex landscape of research in this area, the authors conducted a systematic bibliometric analysis, to rigorously explore and analyse the scientific literature [18]. By delving into the historical evolution, current status, and emerging trends within privacy protection research in AI, this paper aims to provide valuable insights that can guide future research endeavors and policy initiatives.

Insights from comprehensive bibliometric analyses align with these policy and regulation demands. By mapping out the major themes and identifying gaps in AI privacy research, such studies can offer guidance to policymakers. They provide a foundation for developing informed regulations that address both current and future privacy concerns in AI, demonstrating the direct link between academic research and practical policy-making initiatives. Moreover, bibliometric analyses can steer future research. By identifying areas in AI privacy that have been under-explored, these analyses can guide future research efforts toward these priorities, ensuring that academic work is in harmony with national policy objectives and contributes to tackling the most urgent challenges in AI privacy. Additionally, by pinpointing the methodologies and solutions frequently discussed for privacy protection in AI, bibliometric studies can aid developers in integrating best practices into their work. Such analyses offer a resource for developers by spotlighting effective privacy protection methods that comply with regulatory standards.

Extensive bibliometric analyses have been conducted across diverse AI-related domains, including agriculture [19], education [20], [21], energy systems [22], healthcare [23], marketing [24], the maritime industry [25], engineering [26], [27], finance [28], and accounting and auditing [29].

Whereas some bibliometric analysis articles touch on aspects related to privacy protection in AI environments, they exhibit certain limitations in fully addressing this research gap. For instance, some of these articles are primarily focused on specific situations, such as the context of the COVID-19 pandemic [30]. Others, although related, emphasize the ethics dimension of AI [31], [32], [33] or conduct comparative analyses in specific areas [34]. It is the authors' opinion that these limitations highlight the need for a dedicated and systematic bibliometric analysis specifically centered on privacy protection within AI environments. In light of this omission in scholarly literature, this research paper endeavors to fill this gap. By meticulously examining the extensive body of knowledge in this field, the authors aim to provide an all-encompassing perspective on the research landscape and make a contribution to the ongoing discourse surrounding the preservation of privacy in the era of AI.

II. METHODS

The objective of this study is to evaluate previous and ongoing scholarly work on privacy preservation in the context of AI using a bibliometric approach. This paper aims to answer the following questions:

- What are the dominant research trends in AI-related privacy protection?
- What are the main patterns in the distribution of topics (keywords) in the field of AI and privacy protection in past studies?
- Which journals and papers play a prominent role in this research?
- Who are the leading countries and authors contributing to this field?
- What research gaps and challenges can be identified within AI privacy protection research? Are there emerging subfields or niche areas?
- What recommendations can enhance future research on preserving privacy in AI environments?

In pursuit of this objective, the authors implemented the following procedure for conducting the bibliometric analysis, as depicted in the illustrative flowchart presented in Figure 1.

A. DATA ACQUISITION

Initially, the authors compared the most popular bibliographic databases, namely Web of Science (WoS) and Scopus [35]. The authors employed a preliminary set of fundamental search terms relevant to the research topic, such as (''AI'' OR ''Artificial Intelligen*'') AND (''Privacy'' OR ''Person* Data'' OR ''Data Privacy''). Subsequently, they retrieved data from these databases and performed initial analyses on each set of search results. These analyses aimed to evaluate the pertinence of the retrieved data to the research topic and assess the data's overall quality. After careful examination, WoS emerged as the preferred choice due to its superior data quality, which included fewer missing fields and more standardized values. Consequently, the authors chose to utilize the Web of Science Core Collection database, which includes all available editions. Considering the interdisciplinary nature of this research, the authors formulated a systematic search strategy by amalgamating an established AI search strategy from a previous study [36] with a newly devised approach tailored to privacy protection. The

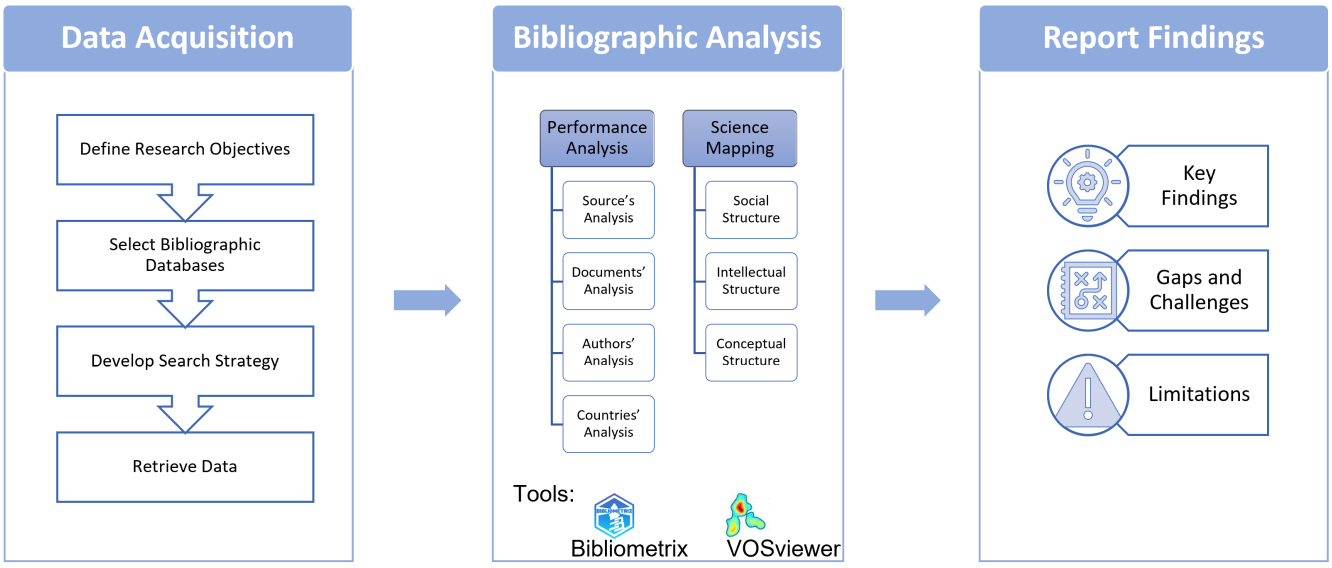


FIGURE 1. Procedure adopted for bibliometric analysis.

process of crafting the privacy protection search strategy was iterative, commencing with the initial search term "privacy protection" to retrieve all articles related to privacy protection and subsequently assessing the frequency of keywords within the dataset. Keywords related to privacy protection that exhibited high frequencies were then incorporated into the search strategy. Through multiple iterations and refinements, the final search strategy emerged as a fusion of both the AI and privacy protection strategies (see Table 1).

The authors screened search results by limiting them to English language documents, encompassing journal articles, conference proceedings, early access publications, review articles, and book chapters. To ensure the accuracy of yearly productivity calculations and avoid inaccuracies caused by incomplete data for 2024, articles from that year were excluded from the dataset. This process resulted in a dataset comprising 8,322 records.

B. BIBLIOGRAPHIC ANALYSIS

The authors employed bibliometric analysis to analyze the articles retrieved from the Web of Science database.

Bibliometric analysis is a quantitative method for assessing scholarly productivity and identifying trends across disciplines [18]. It was first introduced by Alan Pritchard in 1969 and has since become a valuable tool in the realm of scientometrics, which focuses on the study of scientific publications and their impact [37]. These methods were utilized to investigate the current state of research and emerging trends in the field of AI privacy protection. The applied techniques encompassed performance analysis and science mapping [38].

Performance analysis entails the utilization of bibliographic data as a means to evaluate the productivity and influence of various scientific stakeholders, typically including authors, academic institutions, nations, and scholarly journals [18].

Metrics used for performance analysis including:

- Local Citation (LC): Local citations are the references that an article receives exclusively from other articles within the same review corpus, filtered to include only those within the review domain. When an article garners a significant number of local citations, it indicates its impact and influence within the specific discipline [18].
- h-index: The h-index (Hirsch index) refers to the number of articles published by an author (or a journal) that have received at least h citations each from other articles [39]. It measures both the productivity and citation impact of a scholar's publications, but does not account for highly cited outliers and may not accurately reflect the impact of early-career researchers.
- g-index: The g-index, introduced by Egghe in 2006, is an improvement over the h-index and is used to measure the overall citation performance of a set of articles [40]. To calculate the g-index, the set of articles is ranked in decreasing order based on the number of citations they received. The g-index is then the largest number such that the top g articles together have received at least g^2 citations. It helps to recognize scholars whose work has a broad impact, particularly in cases where a few papers receive an exceptionally high number of citations. It addresses the issue of the h-index not sufficiently rewarding highly-cited papers.
- m-index: The m-index is defined as the h-index (h) divided by the number of years (n) since the first published paper of the scientist (or journal) [41]. It provides a measure of the average impact per year since the researcher's (or journal's) first publication. Higher values of these indices indicate greater impact and influence of the source in the academic community.

TABLE 1. Search terms.

Theme	Search Term
#1 AI	TS = (“Artificial Intelligen” or “Neural Net*”or “achine* Learning” or “Expert System$” or “Natural Language Processing” or “Deep Learning” or “Reinforcement Learning” or “Learning Algorithm$” or “*Supervised Learning” or “Intelligent Agent*”)
#2 AI	TS = ((“Backpropagation Learning” or “Back-propagation Learning” or “Bp Learning”) or (“Backpropagation Algorithm*” or “Back-propagation Algorithm*”) or “Long Short-term Memory” or ((Pcnn$ not Pcnnt) or “Pulse Coupled Neural Net*”) or “Perceptron$” or (“Neuro-evolution” or “Neuroevolution”) or “Liquid State Machine*” or “Deep Belief Net*” or (“Radial Basis Function Net*” or “Rbfnn*” or “Rbf Net*”) or “Deep Net*” or “Autoencoder*” or “Committee Machine*” or “Training Algorithm$” or (“Backpropagation Net*” or “Back-propagation Net*” or “Bp Network*”) or “Q learning” or “Convolution* Net*” or “Actor-critic Algorithm$” or (“Feedforward Net*” or “Feed-Forward Net*”) or “Hopfield Net*” or “Neocognitron*” or “Xgboost*” or “Boltzmann Machine*” or “Activation Function$” or (“Neurodynamic Programming” or “Neuro dynamic Programming”) or “Learning Model*” or (“Neurocomputing” or “Neuro-Computing”) or “Temporal Difference Learning” or “Echo State* Net*”)
#3 AI	TS = (“Transfer Learning” or “Gradient Boosting” or “Adversarial Learning” or “Feature Learning” or “Generative Adversarial Net*” or “Representation Learning” or (“Multiagent Learning” or “Multi-agent Learning”) or “Reservoir Computing” or “Co-training” or (“Pac Learning” or “Probabl* Approximate* Correct Learning”) or “Extreme Learning Machine*” or “Ensemble Learning” or “Machine* Intelligen*” or (“Neuro fuzzy” or “Neurofuzzy”) or “Lazy Learning” or (“Multi* instance Learning” or “Multiinstance Learning”) or (“Multi* task Learning” or “Multitask Learning”) or “Computation* Intelligen*” or “Neural Model*” or (“Multi* label Learning” or “Multilabel Learning”) or “Similarity Learning” or “Statistical Relation* Learning” or “Support* Vector* Regression” or “Manifold Regulari?ation” or “Decision Forest*” or “Generali?ation Error*” or “Transductive Learning” or (“Neurorobotic*” or “Neuro-robotic*”) or “Inductive Logic Programming” or “Natural Language Understanding” or (“Adaboost*” or “Adaptive Boosting”) or “Incremental Learning” or “Random Forest*” or “Metric Learning” or “Neural Gas” or “Grammatical Inference” or “Support* Vector* Machine*” or (“Multi* label Classification” or “Multilabel Classification”) or “Conditional Random Field*” or (“Multi* class Classification” or “Multiclass Classification”) or “Mixture Of Expert*” or “Concept* Drift” or “Genetic Programming” or “String Kernel*” or (“Learning To Rank*” or “Machine-learned Ranking”) or “Boosting Algorithm$” or “Robot* Learning” or “Relevance Vector* Machine*” or “Connectionis*” or (“Multi* Kernel$ Learning” or “Multikernel$ Learning”) or “Graph Learning” or “Naive bayes* Classifi*” or “Rule-based System$” or “Classification Algorithm*” or “Graph* Kernel*” or “Rule* induction” or “Manifold Learning” or “Label Propagation” or “Hypergraph* Learning” or “One class Classifi*” or “Intelligent Algorithm*”)
#4 AI	WC = (“Artificial Intelligence”)
#5 Privacy Protection	TI = (“*Privacy” or “Privacy*” or “personal information” or “*disclosure” or “confidentiality” or “personal data” or “de*identification” or “informed*consent” or “information security” or “data security” or “data protection”)
AI & Privacy Integrated	(#1 OR #2 OR #3 OR #4) AND #5

It is particularly useful for comparing scholars at different stages of their careers, as it normalizes the h-index for career length.

- Single Country Publications (SCP): The SCP indicates the number of publications where all authors belong to the same country, signifying intracountry collaboration [42].
- Multiple Country Publication (MCP): The MCP counts documents with at least one co-author from a different country for each country, measuring international collaboration intensity [42].
- Multiple Country Publication Ratio: The MCP ratio represents the proportion of publications involving multiple country compared to the total number of articles from each country, reflecting the extent of international collaboration.

Meanwhile, the concept of *Science Mapping* is harnessed to visually represent and scrutinize the present as well as the evolving cognitive and social frameworks within a particular domain of research, enables the investigation of scientific knowledge statistically [43]. It provides researchers with a systematic overview of key findings within a research field and tracks the progression of theories and techniques [44].

Commonly employed *Science Mapping* for uncovering hidden patterns include *Social Structure*, *Intellectual Structure*, and *Conceptual Structure*:

- The *Social Structure* reveals the intricate interactions among authors, institutions, and countries within the realm of scientific research. Among various forms of social structures, the *Co-authorship Network* [45] is the most prevalent and widely studied.
- The concept of *Intellectual Structure* elucidates how an author’s contributions impact a particular scientific community. This understanding is facilitated through mechanisms such as *Co-citation Networks* [46] and *Historiographic Mapping* [47].
- *Conceptual Structure* pertains to the fundamental content of scientific discourse, encompassing primary themes and prevailing patterns. It entails the interconnectedness among concepts or terms found within a body of publications.

One method employed for analyzing *Conceptual Structure* involves constructing *Co-word Networks*, where words that appear together in documents are linked within a network. This method serves as a valuable approach for comprehending the subjects addressed within a research domain, identifying noteworthy and contemporary matters. Furthermore, it facilitates the examination of the progression of subjects over time.

The *Co-word Networks* can be used to construct a *Thematic Map* that positions clusters into four quadrants based on their centrality and density (see Figure 2) [48]. Starting from the upper-right quadrant and moving in a clockwise

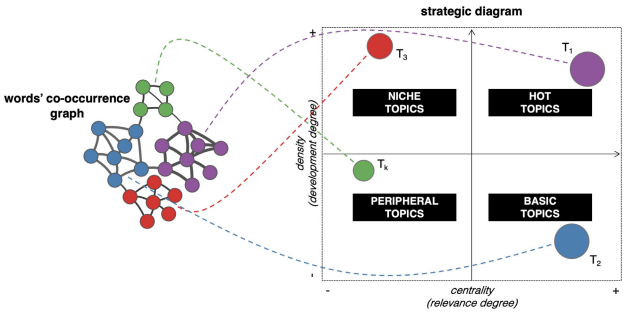


FIGURE 2. Construction of the thematic map from the words' co-occurrence graph [48].

order, the quadrants correspond to motor themes, basic themes, emerging or declining themes, and niche themes, respectively [49].

A *Co-authorship Network* is a collaborative network where individuals are connected based on their shared authorship of academic publications. To construct the co-authorship network in this research, fractional counting is applied for countries with a minimum of five documents. Fractional counting means that when an author co-authors a document with n other authors, each co-authorship link is assigned a strength of $1/n$ [50].

The authors use Bibliometrix R-package [51] and VOSViewer [52] to analyze and visualize the bibliographic data exported from Web of Science.

III. RESULTS

A. CONTEXTUAL OVERVIEW

The dataset includes 8,322 articles in the field of privacy protection in AI environments, covering the period from 1990 to 18th November 2023 (Figure 3). The number of articles has demonstrated substantial growth, indicating an average annual growth rate annual growth rate of 19.4%. This growth has been accelerating since 2018, and 60.5% of the articles in the corpus were published within the last five years.

These articles come from 3,663 different academic or research sources, and they involve a total of 18,346 authors. Within these articles there are 168,305 references to other work, demonstrating the depth of research and knowledge in the field. Additionally, there are 13,256 unique author keywords, highlighting the diverse aspects of privacy protection in AI environments that are being explored. Of the 8,322 articles, 6.6% were single-author documents, with each article having an average of 3.68 co-authors, and an international co-authorship rate of 24.66%, reflecting the high overall level of collaboration in this domain. The analysis also revealed a high level of timeliness and impact, with an average publication age of 5.09 years, and an average of 10.47 citations.

In the 1990s, only a few articles were published on privacy protection in AI, and they received relatively low citation counts. In the early 2000s, the number of articles increased gradually, indicating a growing interest in the topic. In 2002, some articles garnered high interest and gained a lot of citations, such as *k-anonymity: a model for protecting privacy* [53], which accumulated 4,505 total citations. This attracted more researchers to get involved in this area. 2007 saw a small peak in the number of articles, and annual citations continued to increase since the mid to late 2000s. From the early 2010s to 2020, both the number of articles and average citations saw steady growth, reflecting sustained interest in privacy protection in AI (Figure 4). The average citations per article decreased since 2020, primarily because of the fact that some recently published articles need more time to accumulate citations. This decrease may also be partially attributed to the larger volume of research, which could dilute the impact of individual citations.

B. SOURCE ANALYSIS

Analysis of the top 10 sources with the most local impact reveals intriguing patterns (Table 2). *IEEE Transaction on Knowledge and Data Engineering* showcases the highest h-index (36), g-index (68), and total citations (4,921), underscoring its robust local impact in both citations and published papers. *IEEE Access*, with a substantial number of publications (133), signifies its emergence as a pivotal source in the domain. The journal *Data Knowledge Engineering*, commencing research on privacy protection in AI since 1995, stands as one of the earliest contributors, exemplifying its consistent dedication to this field. Although *IEEE Internet of Things Journal* entered this topic in 2018, its notable m-index of 4.833 highlights its strong impact within this arena.

IEEE Transactions on Knowledge and Data Engineering and *Expert Systems with Applications*, which began publishing articles in this specific area in 2002 and 2007, respectively, were pioneering sources in this field, exhibiting a consistent year-over-year increase in published research. *IEEE Access* and *IEEE Internet of Things Journal*, although entering the scene later in 2016 and 2018, respectively, have shown significant growth in the number of published articles, particularly in recent years (Figure 5).

C. DOCUMENT ANALYSIS

1) MOST CITED ARTICLES

In the realm of privacy protection in AI environments, a series of groundbreaking papers have illuminated the path forward for safeguarding sensitive information. Table 3 showcases the articles with the highest global citations. These papers span diverse sub-groups, each contributing to the broader conversation on privacy protection.

a: PRIVACY-PRESERVING TECHNIQUES

One of the foundational works in this domain is Sweeney's *K-Anonymity: A Model for Protecting Privacy* [53]. This paper introduces the concept of k-anonymity, a formal protection model that ensures individuals in shared data cannot be re-identified. With 4,505 citations, it has significantly influenced the field. Abadi et al.'s *Deep Learning with Differential*

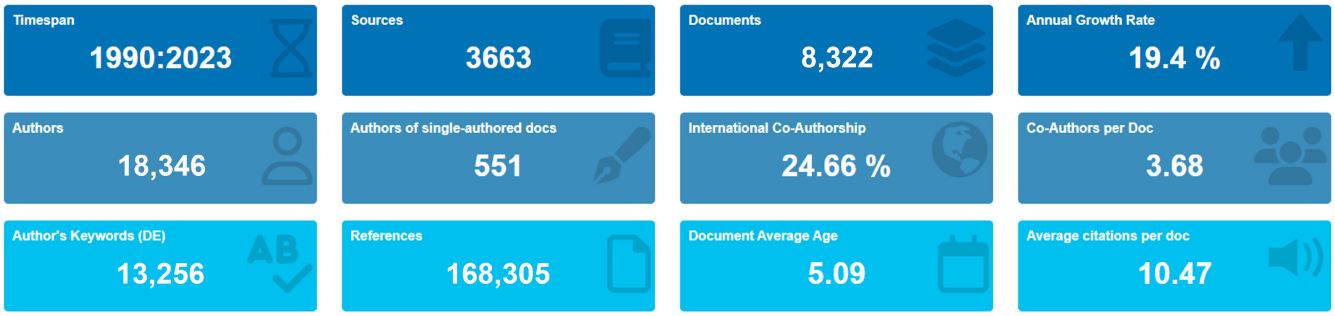


FIGURE 3. Contextual overview.

TABLE 2. Top 10 journals with the highest h-index.

Sources	h-index	g-index	m-index	TC	NP	PY_start
IEEE Transactions on Knowledge and Data Engineering	36	68	1.636	4921	117	2002
IEEE Internet of Things Journal	29	50	4.833	2796	106	2018
Decision Support Systems	26	51	1.238	3405	51	2003
IEEE Access	22	35	2.75	1539	133	2016
IEEE Transactions on Industrial Informatics	22	45	4.4	2126	59	2019
Expert Systems with Applications	20	31	1.176	1154	72	2007
IEEE Transactions on Information Forensics and Security	20	51	1.538	2632	72	2011
Information Sciences	17	29	0.81	871	48	2003
Data & Knowledge Engineering	16	26	0.552	765	43	1995
Knowledge and Information Systems	16	29	0.842	924	43	2005

TABLE 3. Top 10 articles with the highest global citation.

Title	TC	TC/Y	NTC
k-anonymity: a model for protecting privacy [53]	4505	204.77	17.26
Deep Learning with Differential Privacy [54]	1773	221.63	105.31
Practical Secure Aggregation for Privacy-Preserving Machine Learning [55]	1018	145.43	70.66
SecureML: A System for Scalable Privacy-Preserving Machine Learning [56]	724	103.43	50.25
Privacy-Preserving Deep Learning [57]	653	72.56	49.46
Privacy-Preserving Deep Learning via Additively Homomorphic Encryption [58]	604	100.67	38.86
Privacy preserving crowd monitoring: Counting people without people models or tracking [59]	562	35.13	44.08
The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online [60]	507	36.21	29.10
Federated Learning With Differential Privacy: Algorithms and Performance Analysis [61]	487	121.75	39.93
Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT [62]	470	117.50	38.54

Note: TC - Total Citations, TC/Y - Total Citations per Year, NTC - Normalized Total Citations.

Privacy [54] extends the focus to neural networks and sensitive dataset training, providing innovative algorithmic techniques within the framework of differential privacy. Remarkably, it holds the highest total citation per year (221.63). Moving to practical applications, *Practical Secure Aggregation for Privacy-Preserving Machine Learning* [55] presents a secure protocol crucial for federated learning. It ensures efficient and secure aggregation of user-held data vectors. Meanwhile, Mohassel and Zhang's *SecureML: A System for Scalable Privacy-Preserving Machine Learning* [56] addresses privacy concerns in linear and logistic regression, presenting efficient protocols in a two-server model. *Privacy-Preserving Deep Learning* [57], focusing on deep learning, introduces a system allowing multiple parties to jointly learn neural network models without sharing raw datasets. It strikes a balance between utility and privacy. Another paper, *Privacy-Preserving Deep Learning via Additively Homomorphic Encryption* by Phong [58], addresses the challenge of preserving privacy in deep learning and tackles concerns of information leakage. Wei's contribution [61], *Federated Learning With Differential Privacy: Algorithms and Performance Analysis*, introduces a novel framework to prevent information leakage in federated learning, offering valuable insights into privacy-preserving machine learning.

b: PRIVACY IN MONITORING AND INFORMATION DISCLOSURE

Privacy-preserving crowd monitoring: Counting people without people models or tracking [59] employs innovative techniques for estimating crowd sizes without explicit object segmentation, preserving individual privacy. Bansal's work on *The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health*

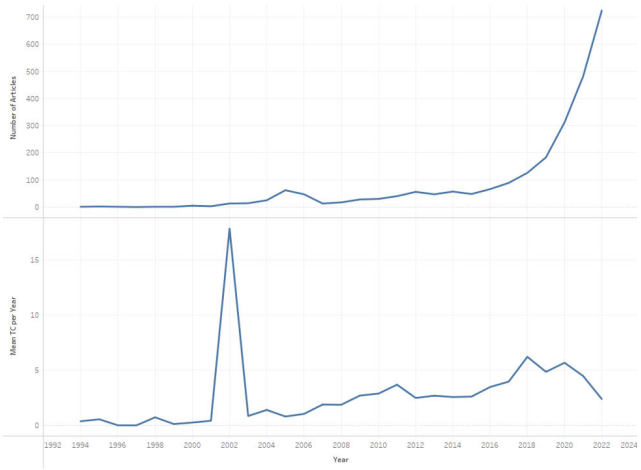


FIGURE 4. Annual scientific production and average citations.

information online [60] delves into the psychology behind disclosing health information, providing insights into building trust in healthcare services.

c: PRIVACY IN INDUSTRIAL IOT AND DATA SHARING

As data sharing becomes increasingly critical, *Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT* by Lu [62] takes center stage, proposing a secure data sharing architecture for Internet of Things (IoT) and emphasizing privacy preservation in industrial applications.

2) MOST FREQUENTLY USED KEYWORDS

The most frequently used indexing keywords, as depicted in the word cloud (Figure 6), reflect a wide spectrum of topics related to privacy protection. In this analysis, these keywords have been organized into a series of distinct themes. Each theme encapsulates specific keywords and highlights important areas of focus within the convergence of privacy and AI research.

The first theme revolves around privacy and security in various contexts. It includes terms related to protecting sensitive data, such as ''differential privacy'', ''homomorphic encryption'', ''k-anonymity'', ''location privacy'', and ''anonymization''. Other keywords highlight techniques for ensuring data security, like ''encryption'', ''cryptography'', ''authentication'', ''confidentiality'', and ''access control''.

The authors have also grouped terms related to machine learning and AI into a theme that covers a wide range of topics, from fundamental concepts such as ''machine learning'', ''training'', ''feature extraction'', ''neural network'', ''classification'', ''privacy-preserving machine learning'', and ''deep learning'' to advanced techniques such as ''federated learning'', ''Convolutional Neural Network'', ''reinforcement learning''. It also includes terms related to specific AI applications, such as ''natural language processing'', ''face recognition'', and ''generative adversarial networks''.

The third group focuses on emerging technologies that are shaping the digital landscape. Terms like ''Internet of Things'', ''cloud computing'' and ''edge computing'' highlight the role of connectivity and data processing in modern systems. ''Blockchain'' and ''smart contract'' point towards decentralized and secure transaction methods. ''Big data'', ''personal data'' and ''data sharing'' underline the importance of data in these technological advancements. This group also includes terms involving urban development and energy management, such as ''smart city'' and ''smart grids'', illustrating the integration of intelligent technologies for efficient and sustainable urban living.

The fourth theme centers on data analysis and modeling. Terms like ''computational modeling'' and ''data models'' emphasize the use of computational techniques for understanding complex systems. ''Data mining'', ''privacy preserving data mining'', and ''clustering'' highlight the extraction of patterns from large datasets, whereas ''task analysis'' and ''optimization'' underscore the goal of improving processes and decision-making.

Finally, the fifth theme pertains to collaboration and social aspects of computer science. Terms like ''protocols'', ''secure multiparty computation'' signify the importance of communication and coordination in technological systems. ''Social media'', ''online social networks'', ''GDPR'', and ''privacy policy'' highlight the integration of digital platforms into everyday life, emphasizing the need for secure communication, data protection, and compliance with privacy regulations. Additionally, in the context of recent global events, terms like ''COVID-19'' and ''healthcare'' reflect the evolving landscape and the role of AI in addressing challenges related to health and safety.

These keywords showcase the multifaceted nature of privacy protection research in AI environments, encompassing privacy, machine learning, emerging technologies, data analysis, and collaboration.

D. AUTHOR ANALYSIS

The top 10 authors with the highest h-index in the dataset offer a glimpse into the impactful contributions made in various fields of privacy protection in AI environments. Table 4 displays key metrics for the top 10 authors with the highest h-index. Among them, Jin Li stands out for his highest contribution, with 50 papers and the highest h-index (18), g-index (44), and total citations (1,964) in the dataset. This signifies high impact from well-cited papers in fields like trust and dependable AI, cloud computing, and blockchain. Li's work includes innovations such as secure fuzzy keyword search and reliable key management for cloud deduplication.

Elisa Bertino, among the top 10 authors, has been actively contributing since 1998. Her research scope covers insider threat protection, IoT security, embedded systems, drones, digital identity, cloud data security and privacy, mobile device privacy, and data trustworthiness.

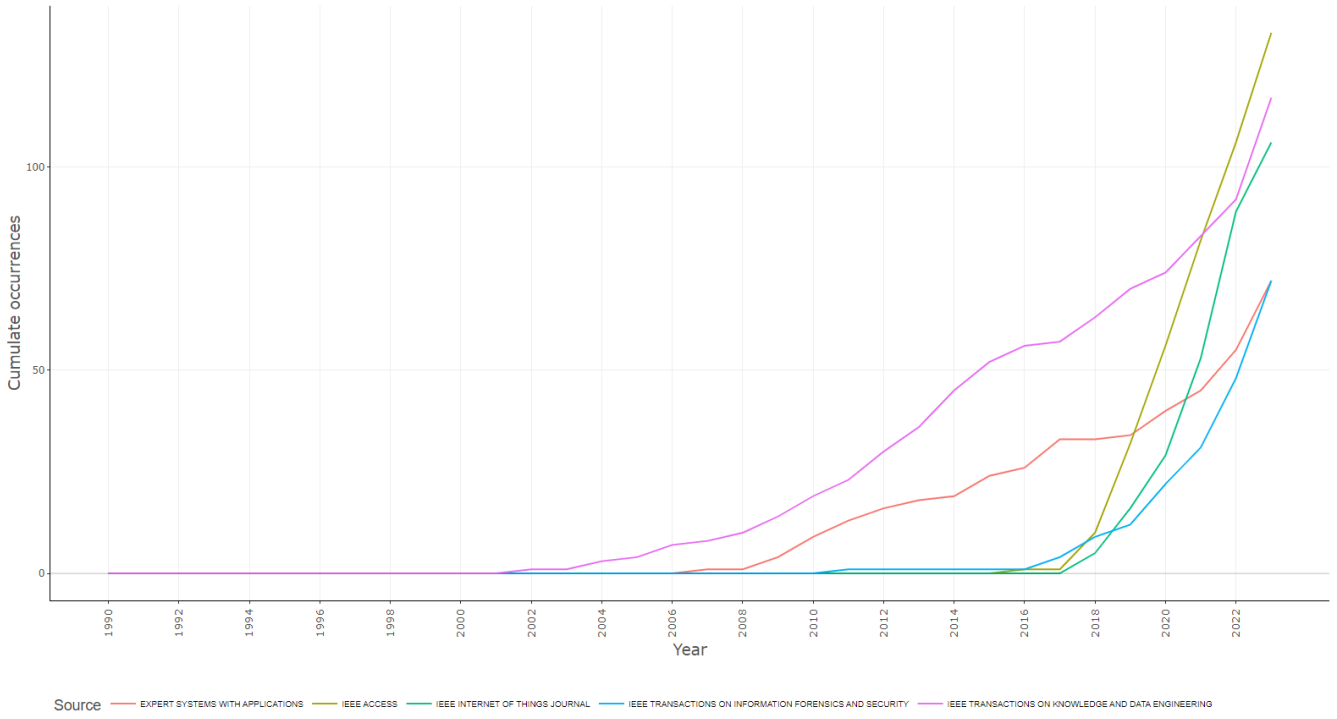


FIGURE 5. Source's production over time.

TABLE 4. Top 10 authors with the highest h-index.

Author	h-index	g-index	m-index	TC	NP	PY_start
Li J	18	44	1	1964	50	2006
Liu XM	16	27	1.455	773	36	2013
Zhang Y	14	34	1.273	1223	40	2013
Polat H	13	22	0.619	509	24	2003
Bertino E	12	22	0.462	561	22	1998
Choo KKR	12	18	2	355	25	2018
Li P	12	25	1.714	970	25	2017
Yu S	12	23	1.091	549	26	2013
Domingo-Ferrer J	11	28	0.5	836	28	2002
Li HW	11	14	2.2	548	14	2019

E. AFFILIATION ANALYSIS

The top 10 most relevant affiliations, according to the number of articles published, are located in China, the US, and Australia, respectively (Table 5). The Chinese Academy of Sciences has the highest number of publications with 228 articles, followed by Xidian University and the University of California system with 189 and 159 articles, respectively.

Affiliations in China with the highest number of articles are the Chinese Academy of Sciences, Xidian University, University of Electronic Science and Technology of China, and Tsinghua University, which together contributed a total of 600 articles. The top affiliations in the US are the University of California system, University of Texas system, Pennsylvania Commonwealth System of Higher Education (PCSHE), Carnegie Mellon University, and State University System of Florida, which together contributed a total of 593 articles.

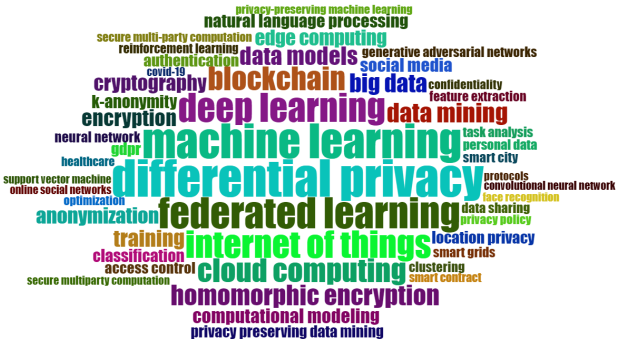


FIGURE 6. Word cloud - the top 50 most frequently used indexing keywords.

The leading affiliation in Australia is the University of Technology Sydney, which contributed 101 articles.

TABLE 5. Top 10 most relevant affiliations.

Affiliation	Articles	Country
Chinese Academy of Sciences	228	China
Xidian University	189	China
University of California System	159	US
University of Texas System	142	US
Pennsylvania Commonwealth System of Higher Education (PCSHE)	103	US
University of Technology Sydney	101	Australia
Carnegie Mellon University	95	US
State University System of Florida	94	US
University of Electronic Science and Technology of China	93	China
Tsinghua University	90	China

Note: This table lists the top affiliations based on the number of articles published.

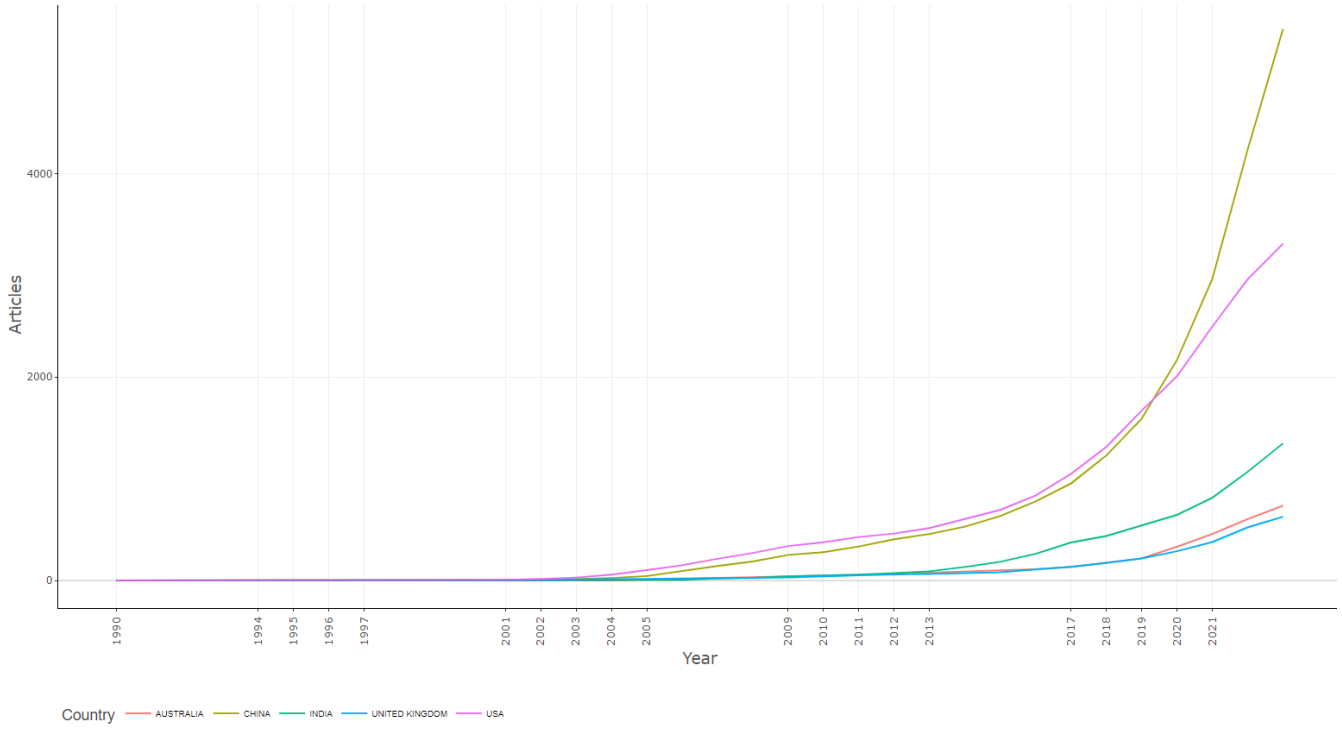


FIGURE 7. Country production over time.

F. COUNTRY ANALYSIS

The USA was among the earliest participants in this research domain and established a leading position, with a steady and substantial increase starting from the mid-2000s. Notably, the growth rate accelerated, particularly during the period from 2020 to 2023.

China, whereas entering the scene later than the USA, has displayed a remarkable ascent since the early 2000s. It experienced exponential growth in scientific output, ultimately surpassing other nations. Distinct spikes in article production emerged from 2016 onwards, enabling China to overtake the USA and claim the top spot in 2019 (Figure 7). This trend further escalated with an exceptional surge between 2020 and 2023.

Australia and the United Kingdom have also exhibited positive trends in scientific production, marked by consistent annual increases in the number of published articles.

India, despite commencing with modest article numbers, has steadily elevated its output over time. Notably accelerated growth can be observed from around 2015 onwards. By 2009 and 2011, it surpassed the UK and Australia, respectively, securing the third-place position. A notable increase in articles occurred between 2021 and 2023.

Whereas China and the USA stand out as major players, some regions, particularly in Africa and South America, display comparatively lower scientific production compared to other continents (Figure 8). The map provides a global overview of scientific research activities, unveiling

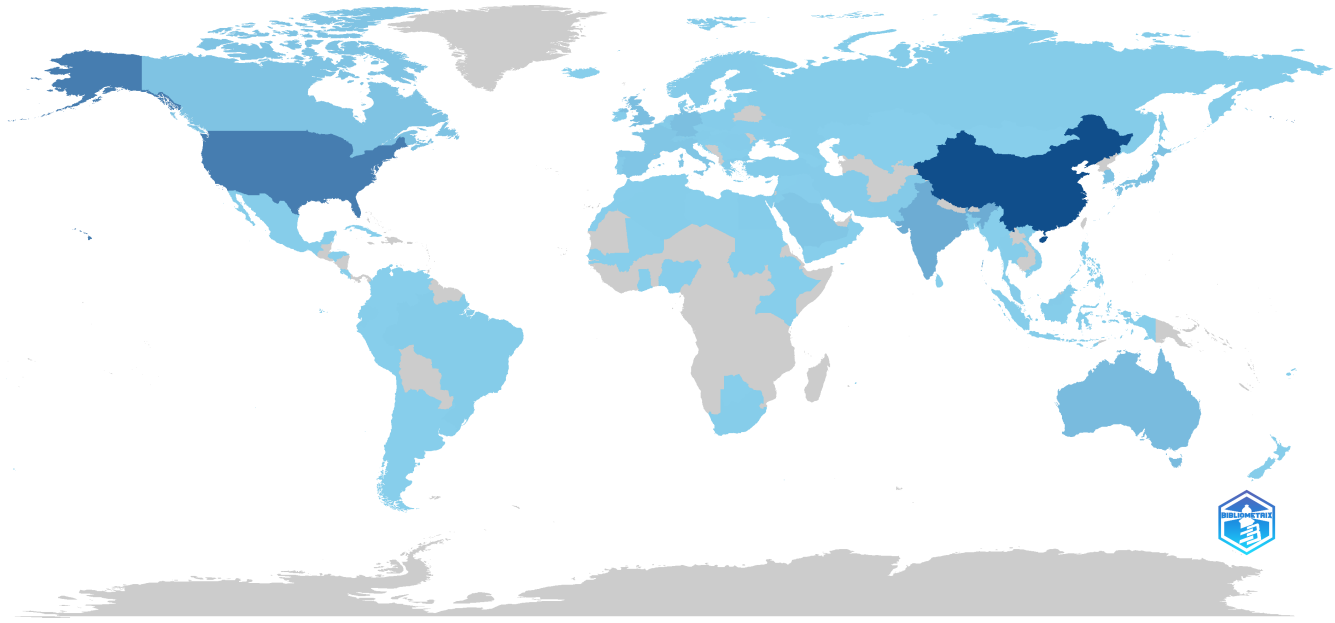


FIGURE 8. Country scientific production - illustrates the scientific production of different countries or regions, representing the frequency of articles contributed by each. The color intensity corresponds to the number of publications, with the gray area indicating no values.

substantial disparities in scientific productivity among countries and regions.

G. SOCIAL STRUCTURE

1) CO-AUTHORSHIP NETWORK

The co-authorship network offers insights into the social interactions and relationships among countries and regions, illustrating their contributions to the advancement of the research field [63]. Out of 114 countries in the dataset, 62 with a minimum of 15 articles are included in the analysis. These countries are clustered into five groups, each distinguished by different colors (Figure 9). Analyzing the co-authorship map has revealed several findings:

a: CENTRAL GLOBAL PLAYERS

The light blue cluster, centrally positioned on the map, represents major global players like China, the United States, Australia, and Japan. These countries exhibit substantial collaboration strength, with China notably displaying high weights across various collaboration metrics. This underscores their significant presence and influence on a global scale.

b: GEOGRAPHICAL PROXIMITY COLLABORATION

A clear trend is observed where countries tend to cluster based on their geographical proximity. For instance, European countries such as Germany, Spain, Italy, Netherlands, Switzerland, Austria, Sweden, Greece, Portugal, and others form a distinct and tight-knit red cluster in the top left. This suggests robust research ties within Europe. Additionally, a country's position within a cluster provides insights into its collaboration intensity, with closer proximity to the cluster center indicating higher collaboration levels, as exemplified by Italy in the red cluster.

A similar strong regional focus is observed in the purple cluster situated in the top right. Comprising countries/regions primarily from Asia such as Singapore, South Korea, and Taiwan, this cluster highlights active collaboration within the Asian research community.

Conversely, the green cluster in the bottom right includes countries primarily from South Asia and the Middle East, including India, Saudi Arabia, Algeria, Egypt, Jordan, Qatar, Tunisia, and the United Arab Emirates, emphasizing their active participation and engagement in collaborative research efforts within the region.

c: DIVERSE COLLABORATION PATTERNS

The dark blue cluster demonstrates diverse collaboration patterns, transcending multiple regions. Canada, for instance, has strong collaborations with China and the United States, as well as diverse connections with countries in different regions. On the other hand, England, although closer to the red cluster of European countries, has dense trans-regional connections with other countries around the world.

2) COUNTRIES' SCIENTIFIC PRODUCTION AND COLLABORATION

The scientific production demonstrates a significant pattern, with the top 10 countries contributing 74.8% of the total publications. In particular, China and the USA together account for 46.1% of this production, highlighting their prominent roles (Table 6).

The USA has the highest Total Citations (TC) at 34,025 and the highest Average Article Citations of 22.90, showcasing its global research impact. Notably, Canada excels with the second-highest Average Article Citations (15.60), indicating the impact of its research within the scientific community. In contrast, China, despite its highest publication

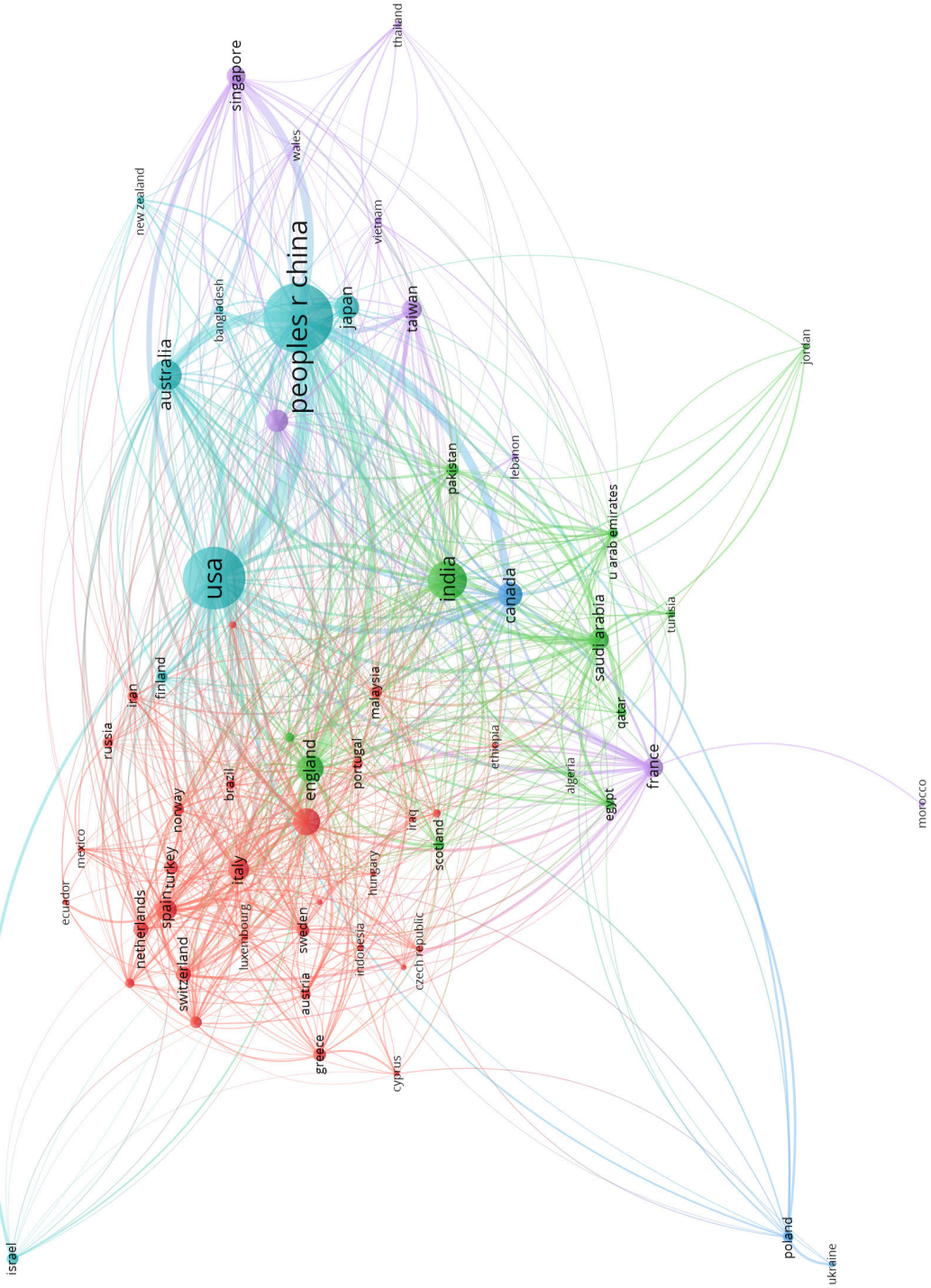


FIGURE 9. Global co-authorship network - the size of the circles corresponds to the fractional counting of the number of articles associated with each country, while the strength and distance of the links represent the number of co-authorships and the relatedness of the articles in terms of co-authorship links.

count (2,352), exhibits a lower Average Article Citations of 8.40, suggesting potential for improving the quality and impact of individual papers.

Australia demonstrates the highest MCP Ratio (0.397), emphasizing a strong focus on international collaborative research efforts. It is followed by Canada (0.363), the UK (0.356), and China (0.268), all illustrating strong international collaboration in this field. The USA and India, though ranking second and third in terms of publications, have the lowest MCP ratios among the top 10 contributing countries, with MCP ratios of 0.143 and 0.112, respectively. This suggests an emphasis on intra-country collaboration rather than international collaboration.

H. INTELLECTUAL STRUCTURE

The historiographic analysis uncovers the evolutionary journey of privacy protection within AI environments. It reveals research paths evolving over time (see Figure 10 and Table 7).

1) CLUSTER 1: ALGORITHM-ORIENTED PRIVACY PROTECTION

Cluster 1, which commenced in 2002, exhibits an emphasis on algorithm-oriented approaches. Commencing with Sweeney's k-anonymity model, these papers introduce methodologies to safeguard privacy by transforming data into formats conducive to analysis while safeguarding individual identities. This cluster encompasses techniques such as k-anonymity and de-identification, striving to harmonize data utility with the preservation of privacy. Notably, this cluster charts a progression from initial models like k-anonymity to more advanced algorithms focused on de-identifying facial images, signifying an ongoing exploration of algorithmic privacy solutions.

2) CLUSTER 2: DATA-ORIENTED PRIVACY PRESERVATION AND BLOCKCHAIN

Cluster 2, originating in 2004, revolves around data-oriented privacy preservation techniques. These studies concentrate on scenarios where data is horizontally partitioned across various entities, leading to a pronounced emphasis on secure distributed data mining. Topics covered encompass cryptographic methods, secure multi-party computation, and integration of blockchain-based encryption to facilitate data sharing while upholding privacy. This cluster signifies an evolution from early privacy-preserving data mining tactics to intricate models involving the integration of blockchain technology.

3) CLUSTER 3: CLOUD-COMPUTING ORIENTED PRIVACY SOLUTIONS

Initiating from 2014, Cluster 3 underscores cloud-computing oriented privacy solutions. These papers delve into harnessing cloud resources for privacy-preserving computations, often entailing encryption and secure computation techniques. This topic encompasses operations performed on encrypted data within cloud environments, signifying a shift towards leveraging cloud computing capabilities while safeguarding data privacy. This cluster reflects the progressive evolution of privacy-preserving solutions within cloud-based settings.

4) CLUSTER 4: FEDERATED LEARNING AND DIFFERENTIAL PRIVACY

Cluster 4, emerging in 2016, revolves around federated learning and differential privacy. These papers spotlight the burgeoning trend of privacy-preserving collaborative learning models. Encompassing themes of federated learning, differential privacy, and secure aggregation, this cluster illustrates a progression from initial explorations to more sophisticated methodologies. This cluster epitomizes the present-day emphasis on federated learning and differential privacy for training models across decentralized participants.

In summary, this historiograph offers a window into the evolving focus of privacy protection within the AI landscape. The distinct clusters trace a trajectory from algorithm-centric strategies to more intricate data-centric methodologies, evolving further into cloud-computing solutions, and culminating in the contemporary prominence of federated learning and differential privacy. This evolution underscores the field's dynamic response to emerging challenges and transformative technologies in the realm of privacy-preserving AI.

I. CONCEPTUAL STRUCTURE

1) CO-WORD ANALYSIS

The co-word analysis unveils patterns of co-occurrence among keywords, shedding light on relationships between ideas within the subject areas discussed in the texts [80]. The analysis identified and color-coded five clusters, each capturing a cohesive set of research topics within the field of privacy protection in AI environments (Figure 11):

Red Cluster (Emerging Technologies and Distributed Systems):

The red cluster primarily focuses on emerging technologies and distributed systems, encompassing research on blockchains, Internet of Things (IoT), edge computing, distributed computing, reinforcement learning, federated learning, collaborative work, fog computing, wireless computational modeling, data models, servers, training, protocols, etc. Publications in this cluster tend to have a relatively recent average publication year, mostly after 2021, with an increasing trend in the number of publications each year. This suggests the timeliness of research addressing privacy challenges in emerging technologies and communication.

Green Cluster (Human Factors and Compliance):

The green cluster centers around human factors and compliance, covering topics such as trust, risk assessment, risk management, access control, disclosure, data sharing, social networks, decision-making, personalization, ethics, fairness, awareness, e-commerce, GDPR, social networks, and social media, among others. This cluster has the lowest average occurrence, suggesting that its terms are

TABLE 6. Countries' scientific production and collaboration for the top 10 countries.

Country	Articles	SCP	MCP	Freq	MCP Ratio	TC	Avg. Citations
China	2352	1721	631	0.283	0.268	19659	8.40
USA	1483	1271	212	0.178	0.143	34025	22.90
India	672	597	75	0.081	0.112	2680	4.00
Australia	292	176	116	0.035	0.397	3243	11.10
Germany	278	219	59	0.033	0.212	2087	7.50
UK	261	168	93	0.031	0.356	2182	8.40
Japan	249	208	41	0.03	0.165	1768	7.10
Korea	237	178	59	0.028	0.249	1915	8.10
Canada	204	130	74	0.025	0.363	3189	15.60
Spain	200	153	47	0.024	0.235	2089	10.40

Note: The "Freq" column indicates the proportion of articles contributed by each country compared to the total dataset.

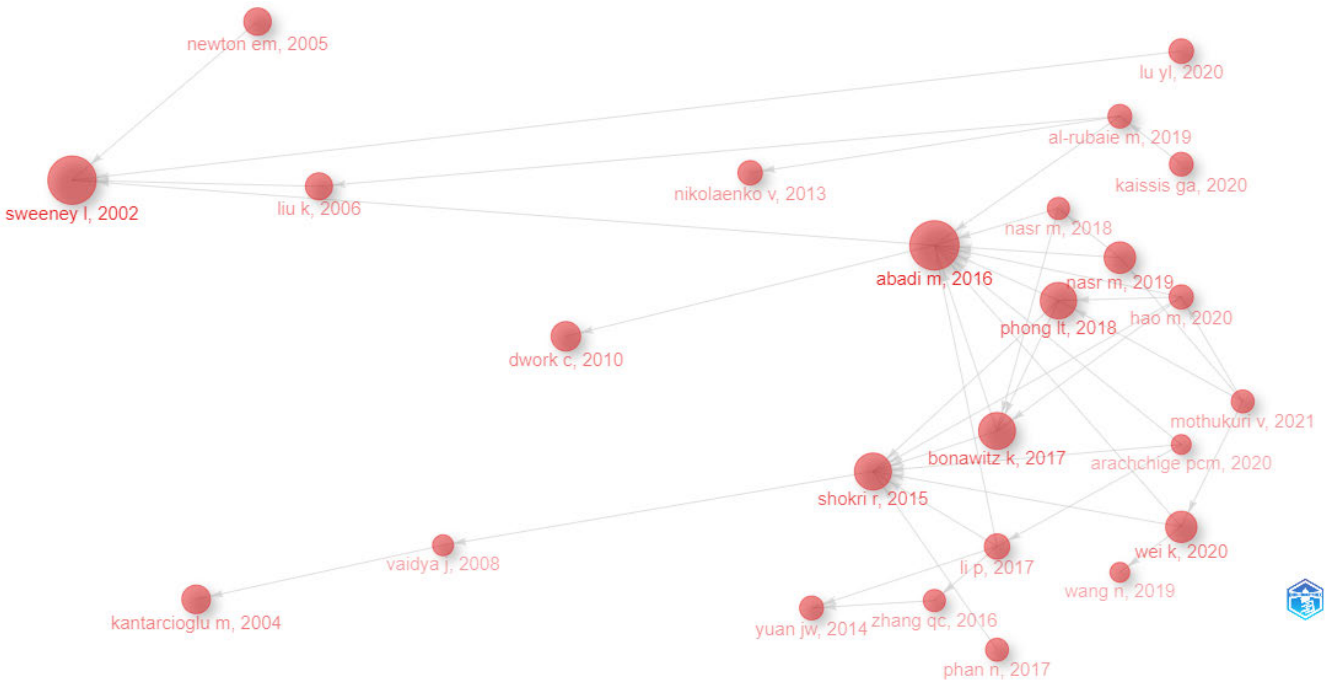


FIGURE 10. Historiograph - illustrating document citations and publication years, offering insights into research trend evolution. Nodes represent cited documents, and edges show direct citations.

specialized and specific to a narrow context. These terms have limited connections with the broader vocabulary, indicating a fragmented and less coherent theme. Compared to other clusters, keywords in these areas exhibit the lowest average links, lowest average link strength, and lowest average citations. This suggests that these topics may be relatively underexplored. Additionally, the average publication year for this cluster is earlier than that of other clusters, with an average publication year of 2018. This suggests that the cluster has fewer recent publications compared to other clusters, implying that research in human factors and compliance in privacy protection in the AI environment is still a less mature or less established research area within the analyzed dataset.

Blue Cluster (Machine Learning and Data Analysis):

The blue cluster revolves around machine learning, data analysis, and big data analytics, incorporating terms related to differential privacy, data mining, adversarial training, anomaly detection, Convolutional Neural Network (CNN), decition tree, deep learning, Deep Neural Network (DNN), Generative Adversarial Network (GAN), neural network, collaborative learning, computer vision, classification, feature selection, support vector machine, and random forest, among others. Publications within this cluster span a wide range of years, reflecting both established and emerging research areas in Machine Learning and Data Analysis.

Yellow Cluster (Cryptography and Network Security): The yellow cluster focuses on cryptographic techniques, encryption, blockchain technology, and topics related to network security, encompassing blockchain, cloud computing, cryptography, data security, encryption, security, confidentiality, authentication, location privacy, smart city, and smart home. Terms in this cluster have the highest average occurrence and citation, implying that topics

TABLE 7. Paper details for historiograph.

Paper	Title	LCS	GCS	Cluster
SWEENEY L, 2002	K-Anonymity: A Model for Protecting Privacy [53]	279	4505	1
KANTARCIOGLU M, 2004	Privacy-Preserving Distributed Mining of Association Rules on Horizontally Partitioned Data [64]	52	364	2
NEWTON EM, 2005	Preserving Privacy by De-identifying Face Images [65]	38	290	1
LIU K, 2006	Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining [66]	40	286	1
VAIDYA J, 2008	Privacy-Preserving Naive Bayes Classification [67]	31	109	1
NIKOLAENKO V, 2013	Privacy-Preserving Ridge Regression on Hundreds of Millions of Records [68]	58	222	1
YUAN JW, 2014	Privacy Preserving Back-Propagation Neural Network Learning Made Practical with Cloud Computing [69]	34	135	3
SHOKRI R, 2015	Privacy-Preserving Deep Learning [57]	249	653	1
ZHANG QC, 2016	Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning [70]	47	164	3
ABADI M, 2016	Deep Learning with Differential Privacy [54]	647	1773	4
BONAWITZ K, 2017	Practical Secure Aggregation for Privacy-Preserving Machine Learning [55]	245	1018	1
LI P, 2017	Multi-Key Privacy-Preserving Deep Learning in Cloud Computing [71]	65	304	3
PHAN N, 2017	Adaptive Laplace Mechanism: Differential Privacy Preservation in Deep Learning [72]	50	90	4
PHONG LT, 2018	Privacy-Preserving Deep Learning via Additively Homomorphic Encryption [58]	234	604	3
NASR M, 2018	Machine Learning with Membership Privacy Using Adversarial Regularization [73]	47	150	1
AL-RUBAIE M, 2019	Privacy-Preserving Machine Learning: Threats and Solutions [74]	54	125	2
WANG N, 2019	Collecting and Analyzing Multidimensional Data with Local Differential Privacy [75]	43	138	4
NASR M, 2019	Comprehensive Privacy Analysis of Deep Learning Passive and Active White-Box Inference Attacks Against Centralized and Federated Learning [76]	138	468	4
ARACHCHIGE PCM, 2020	Local Differential Privacy for Deep Learning [77]	43	83	4
WEI K, 2020	Federated Learning with Differential Privacy: Algorithms and Performance Analysis [61]	135	487	4
KAISSIS GA, 2020	Secure, Privacy-Preserving and Federated Machine Learning in Medical Imaging [12]	54	284	2
HAO M, 2020	Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence [78]	41	234	4
LU YL, 2020	Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT [62]	59	470	2
MOTHUKURI V, 2021	A Survey on Security and Privacy of Federated Learning [79]	50	331	4

in this cluster are more central, influential, and widely recognized.

Purple Cluster (Privacy-Preserving Techniques):

This cluster centers around privacy-preserving methods and techniques, encompassing areas such as data anonymization, perturbation methods, privacy-preserving data mining, data publishing, data utility, k-anonymity, local differential privacy, genetic algorithms, collaborative filtering, and more.

2) THEMATIC MAP

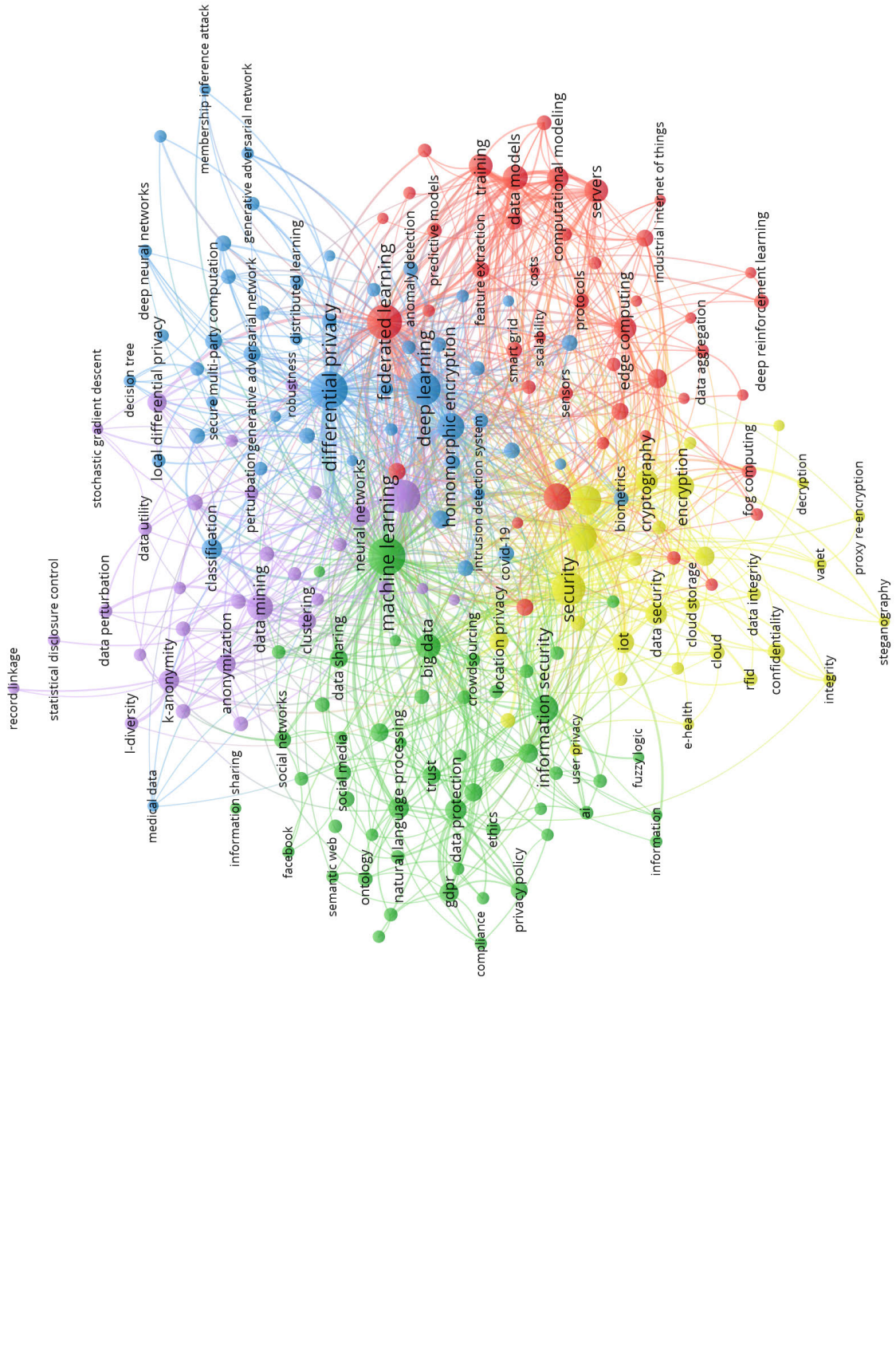
The Thematic Map, derived from the Co-word Networks (as shown in Figure 2) [48], offers a strategic approach to categorizing research themes based on their centrality and density within the network. The resulting quadrants in the thematic map are instrumental in guiding the analysis of the research landscape in the field of study. These quadrants represent distinct types of themes, each signifying different characteristics and implications in the context of the research domain [49] (Figure 12).

Upper-Right Quadrant (Hot Topics / Motor Themes): This quadrant represents themes characterized by high centrality and high density. It includes topics like differential privacy, federated learning, deep learning, Internet of Things, cloud computing, blockchain, homomorphic encryption, data models, training, encryption, cryptography, edge computing, computational modeling, neural network, reinforcement learning, among others. These themes are considered hot topics or motor themes, indicating that they are well-developed and crucial within the research field. They play a central role in structuring the conceptual framework of the domain and hold substantial relevance.

Lower-Right Quadrant (Basic and Transversal Themes): Themes falling in this quadrant exhibit high centrality but low density. Examples include topics like machine learning, data mining, big data, anonymization, classification, authentication, privacy preserving data mining, access control, smart city, etc. These are recognized as basic and transversal themes. Such themes hold importance for the domain and extend across various research areas. They represent overarching concepts that connect different segments of the field.

Lower-Left Quadrant (Emerging or Declining Themes): This quadrant encompasses themes with both low centrality and low density. It includes topics like k-anonymity and location privacy. Themes in this quadrant are emerging or declining, signifying that they are weakly developed and sit at the periphery of research interests. These themes may be in the early stages of exploration or experiencing reduced relevance.

Upper-Left Quadrant (Highly Developed and Isolated Themes): Themes positioned in this quadrant possess well-developed internal links (high density) but lack significant external connections (low centrality). Examples in this quadrant include social media, Natural Language



VOSviewer

FIGURE 11. Co-occurrence analyses for indexing keywords.

Processing (NLP), GDPR, personal data, privacy policy. These are classified as highly developed and isolated themes. For example, critical privacy concepts such as ''GDPR'', ''personal data'', and ''privacy policy'' may not be adequately interconnected with established AI technologies, highlighting a potential gap in interdisciplinary collaboration. This suggests a need for greater synergy between the evolved privacy policies and AI technologies to improve privacy safeguards.

3) RESEARCH TRENDS

The trend topics derived from the frequency of authors' keywords in papers related to privacy in AI environments, as shown in Figure 13, highlight distinctive patterns and shifts in research emphasis over the years. These trends are categorized into specific time periods, each reflecting a dominant theme or focus in the field.

Early Emphasis on Privacy and Data Control (2006-2016): In the earlier years, topics like ''microaggregation'', ''RFID'', ''record linkage'', ''context awareness'', and ''personal information management'' were prevalent. This indicates a strong emphasis on privacy-preserving techniques, risk management, and methods to control data disclosure.

Data Mining and Data Analysis (2010-2020): The terms ''privacy-preserving data mining'', ''distributed data'', ''mutual authentication'', ''randomization'', ''anonymization', ''location privacy'', and ''statistical disclosure control'' gained prominence during this period, highlighting a shift towards data analysis and pattern extraction. Techniques like ''k-anonymity'' were introduced to safeguard sensitive information. Additionally, with the prevalence of ''e-commerce'' and ''risk management'', the concept of ''disclosure'' also gained more attention and recognition.

Rising Influence of Machine Learning and Cloud Computing (2016-2023): Recent years saw a surge in ''machine learning'', ''deep learning'', ''federated learning'', ''cloud computing'', ''blockchain'', ''support vector machine'', ''recommender systems'', ''accuracy'', ''I-diversity'', ''clustering'', ''classification'', ''neural network'', ''big data'', and ''membership inference attack''. These terms demonstrate a transition towards advanced technologies for data analysis, secure computation, and distributed learning. ''Differential privacy'' emerged as a central concept for preserving privacy in data sharing.

Legal Consideration and Healthcare (2018-2023): With the enforcement of ''GDPR'', legal consideration has gain importance since 2018. The recent inclusion of terms such as ''medical services'', ''internet of medical things'' reflects an intensified focus on leveraging technology within healthcare services and enhancing system performance, particularly in the context of the global COVID-19 pandemic. As the healthcare sector increasingly adopts AI applications to manage and analyze sensitive medical data, the importance of privacy protection has been underscored. Striking a balance between utilizing advanced AI-driven solutions and ensuring robust privacy safeguards has become a paramount concern, given the heightened demand for efficient and accurate healthcare solutions in these challenging times.

Overall, the trend topics suggests an evolution from foundational privacy concepts to advanced machine learning techniques and cloud-based approaches. The increasing focus on legal consideration and healthcare underscores the applicability of these technologies in real-world domains. The timeline progression showcases the adaptation of research interests to the evolving technological landscape.

IV. DISCUSSIONS

Although the concept of AI emerged as early as 1955 [81], the progress of AI research has witnessed fluctuations over the decades. The evolution of the AI field has been likened to the changing seasons: spring, summer, winter, and fall, with the most fruitful ''fall'' being observed in recent years [82]. Specifically, research indicates a significant surge in AI publications since 2013, with the number of publications in the five-year span since 2015 equaling that of the preceding 40 years [33]. More precisely, a majority of AI research, exceeding 50% in volume, has been published between 2017 and 2021 [5].

In comparison, the study underscores that publications addressing privacy protection in AI environments markedly accelerated after 2018, displaying a lag of approximately five years following the general acceleration in AI research. This pattern suggests a responsive concern towards the burgeoning AI-related technologies and the increasing prevalence of AI applications, such as facial recognition technology [83], social media algorithms [84], Voice Assistants [85], smart home devices and IoT [86], and healthcare AI applications [87]. High-profile instances, such as the Facebook-Cambridge Analytica scandal [88], Clearview AI data scraping controversy [89], Amazon Alexa privacy issues [90], significantly underscored the urgency of protecting privacy in AI environments. Extensive media coverage of data breaches, privacy violations, and controversies related to AI has markedly heightened public awareness. This heightened awareness has driven an increased demand for research aimed at developing AI privacy solutions and ensuring accountability. Furthermore, the global COVID-19 pandemic has accentuated common concerns regarding data protection and privacy preservation [6].

Furthermore, regulatory changes and legal emphasis have also drawn attention to research in this area. For instance, the implementation of GDPR in 2018 compelled strict data privacy compliance [91], leading to increased research and funding opportunities for developing AI models in alignment with similar regulatory frameworks [92]. Major regions like North America, East Asia, and Europe are strategically advancing AI, with countries implementing policies to boost its application and development [93]. National and international research grants, such as those provided by the National Science Foundation (NSF) in the USA, have been instrumental in funding privacy-related AI research projects [94].

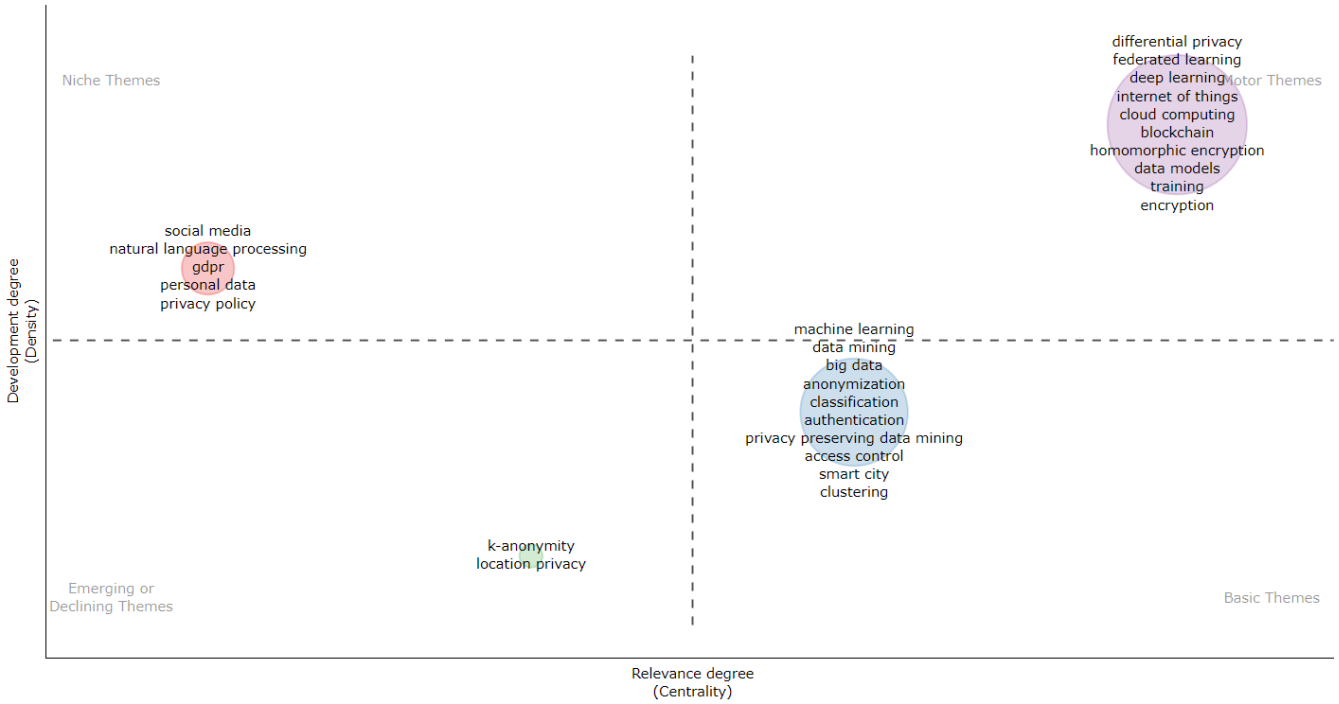


FIGURE 12. Thematic map - the thematic map categorizes research clusters into four quadrants, following a clockwise order from the upper-right quadrant: motor themes, basic themes, emerging or declining themes, and niche themes. Each bubble within the map corresponds to a network cluster, with its size directly proportional to the frequency of word occurrences within that cluster. The label of each bubble represents the word most prominent in the cluster.

A. KEY FINDINGS

The growing concerns about privacy protection in AI environments prompted the authors to analyze past research in this area to answer the following research questions:

1) WHAT ARE THE DOMINANT RESEARCH TRENDS IN AI-RELATED PRIVACY PROTECTION?

Based on the aforementioned analysis, the evolution of research in AI-related privacy protection shows a trajectory shifting from an algorithm-oriented approach to a focus on data orientation, and subsequently to privacy solutions centered around cloud computing. In recent years, there has been a notable shift towards embracing Federated Learning and Differential Privacy.

Early research in this area, represented by Sweeney's k-anonymity model [53], emphasizes algorithm-oriented approaches to safeguard privacy by transforming data into formats conducive to analysis while protecting individual identities. Since 2004, a series of research endeavors, including cryptographic methods, secure multi-party computation, and the integration of blockchain-based encryption to facilitate data mining, sharing, and analysis while upholding privacy, revolve around data-oriented privacy preservation techniques. With the prevalence of cloud computing, cloud-computing oriented privacy solutions gained significance, delving into harnessing cloud resources for privacy-preserving computations, often entailing encryption and secure computation techniques. Since 2016, there is a rising influence of machine learning and blockchain. In recent years, more research focuses on federated learning and differential privacy, spotlighting the burgeoning trend of privacy-preserving collaborative learning models.

2) WHAT ARE THE MAIN PATTERNS IN THE DISTRIBUTION OF TOPICS (KEYWORDS) IN THE FIELD OF AI AND PRIVACY PROTECTION IN PAST STUDIES?

The main patterns in the distribution of topics in the field of AI and privacy protection in past studies reveal a focus on diverse themes, including machine learning and AI, data privacy, emerging technologies and distributed systems, and human factors and compliance, reflecting a comprehensive exploration of fundamental concepts, advanced techniques, connectivity, decentralized systems, and societal implications in the intersection of AI and privacy.

The most prominent keywords in the research area are centered around machine learning and AI, forming a theme that encompasses a wide range of topics. It spans fundamental concepts like ''machine learning'', ''training'', ''feature extraction'', ''neural network'', ''classification'', ''deep learning'', as well as advanced techniques such as ''federated learning'', ''Convolutional Neural Network'', ''reinforcement learning'', and others. Additionally, it includes terms related to specific AI applications, such as ''natural language processing'', ''face recognition'', ''anomaly detection'', and ''generative adversarial networks''.

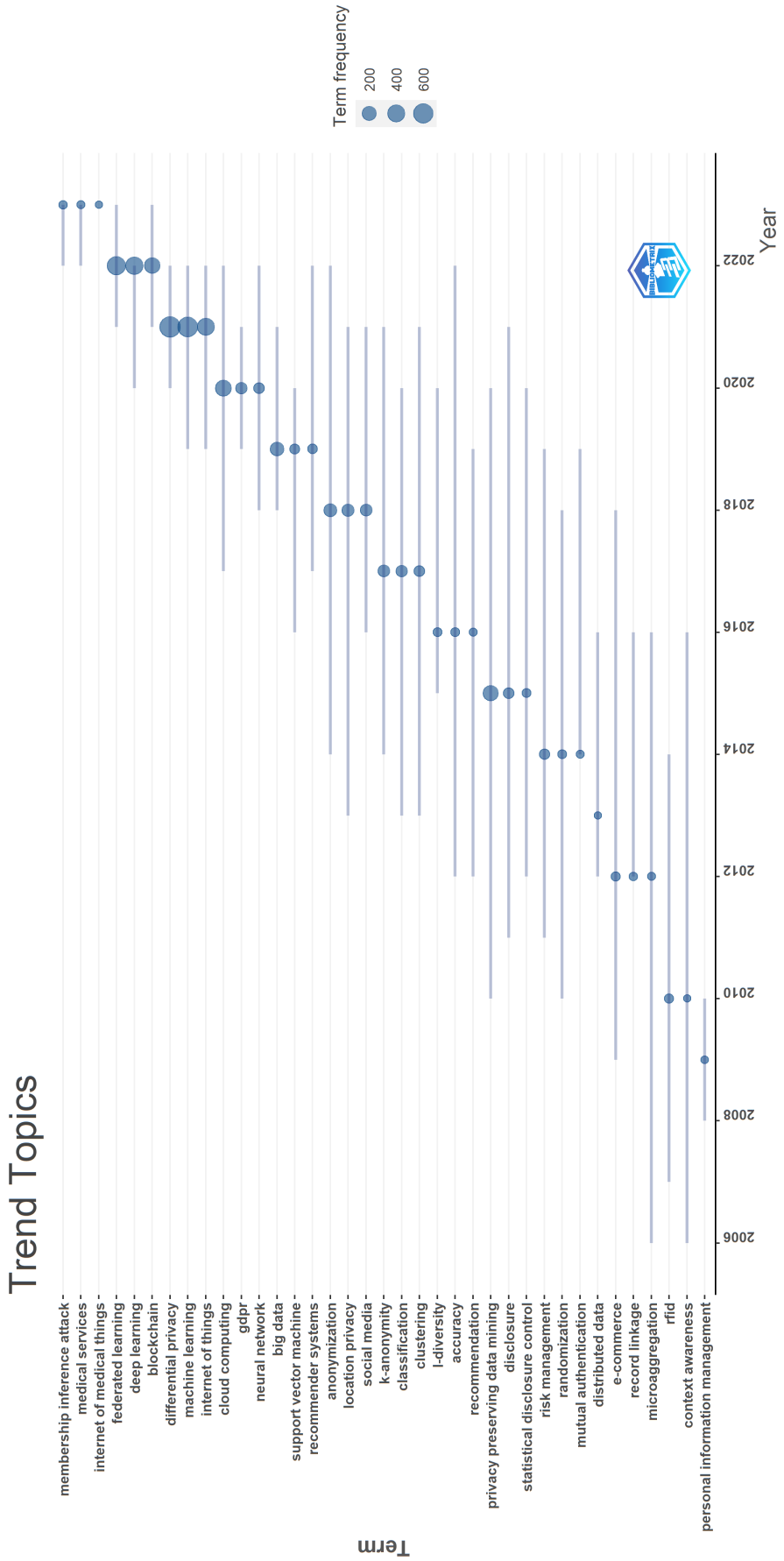


FIGURE 13. Trend Topics - displays the trend topics derived from the frequency of authors' keywords in papers related to privacy in AI environments. The figure illustrates the three most commonly used keywords for each year, with the timeline of articles represented by the line and the bubble sizes proportional to the term frequencies.

Another notable theme revolves around data privacy, incorporating terms related to safeguarding sensitive data. It includes concepts like ''differential privacy'', ''homomorphic encryption'', ''k-anonymity'', ''location privacy'', ''anonymization'', ''big data'', ''personal data'', ''data sharing'', ''data mining'', ''cryptography'', ''authentication'', ''confidentiality'', ''access control'', ''data perturbation'', ''data publishing'', ''data utility'', among others.

Emerging Technologies and Distributed Systems represent another prominent topic. Terms like ''Internet of Things'', ''cloud computing'', and ''edge computing'' highlight the role of connectivity and data processing in modern systems. Additionally, ''blockchain'' and ''smart contract'' indicate the adoption of decentralized and secure transaction methods. This category also includes terms related to urban development and energy management, such as ''smart city'' and ''smart grids'', illustrating the integration of intelligent technologies for efficient and sustainable urban living.

Finally, the fourth theme focuses on Human Factors and Compliance. Terms like ''protocols'' and ''secure multiparty computation'' underscore the importance of communication and coordination in technological systems. Concepts like ''social media'' and ''online social networks'' highlight the integration of digital platforms into everyday life. The theme extends to include topics such as ''GDPR'', ''privacy policy'', ''healthcare'', ''trust'', ''risk assessment'', ''risk management'', ''access control'', ''disclosure'', ''data sharing'', ''decision-making'', ''personalization'', ''ethics'', ''fairness'', ''awareness'', ''e-commerce'', among others.

3) WHICH SOURCES AND PAPERS PLAY A PROMINENT ROLE IN THIS RESEARCH?

The study of privacy protection in AI environments, although initiated in the 1990s, did not garner significant attention until recent years. Approximately 60.5% of the articles in the corpus were published in the last five years, underscoring the burgeoning nature of this academic research field. The journal *Data Knowledge Engineering*, a consistent contributor since 1995, has played a foundational role in the development of the field, showcasing its dedication over the years. *IEEE Transaction on Knowledge and Data Engineering* stands out with the highest h-index, g-index, and total citations, indicating its robust local impact in both citations and published papers. It has been a pioneering source since the early stages of research in this domain, demonstrating a consistent year-over-year increase in published research. Despite entering the scene later, *IEEE Access* has emerged as a pivotal source, exhibiting substantial growth in the number of published articles in recent years.

In the field of privacy protection in AI environments, foundational contributions have paved the way for safeguarding sensitive information. Sweeney's *K-Anonymity: A Model for Protecting Privacy* [53] stands as a seminal work, significantly influencing the field and inspiring researchers to delve into this critical area. Abadi et al.'s *Deep Learning with Differential Privacy* [54], boasting the highest total citation per year, expands the focus to neural networks and sensitive dataset training. This work introduces innovative algorithmic techniques within the framework of differential privacy, contributing significantly to the advancement of research in this realm.

4) WHO ARE THE LEADING COUNTRIES, AFFILIATIONS, AND AUTHORS CONTRIBUTING TO THIS FIELD?

The scientific landscape reveals a distinct pattern, with the top 10 countries—China, the USA, India, Australia, Germany, the UK, Japan, Korea, Canada, and Spain—collectively contributing a substantial 74.8% of total publications. Notably, China and the USA jointly contribute 46.1%, underscoring their pivotal roles.

The USA, an early participant in this research domain, has maintained a leading position, exhibiting consistent growth since the mid-2000s, with an accelerated rate from 2020 to 2023. The USA has the highest total citations and average article citations, affirming its global research influence. China, although entering the field later, has experienced exponential growth since the early 2000s, surpassing other nations and claiming the top spot in 2019. However, China's lower average article citations suggest room for improvement in the quality and impact of individual papers. Whereas China and the USA stand out as major contributors, some regions, particularly in Africa and South America, exhibit comparatively lower scientific production.

Geographical proximity influences collaboration patterns, evident in European, Asian, and Middle Eastern countries. Conversely, countries like Canada and the UK showcase diverse collaboration patterns transcending multiple regions. Among the top productive countries, Australia, Canada, the UK, and China emphasize international collaborative research, whereas the USA and India prioritize intra-country collaboration.

The top 10 affiliations with the most publications are predominantly located in China, the US, and Australia, highlighting a major concentration of research in these countries. Specifically, the Chinese Academy of Science, Xidian University, and the University of California system are at the forefront in terms of publication numbers, with four affiliations from China, five from the US, and one from Australia among the leading contributors.

Authors contribute significantly to the field through either highly cited single papers or a high h-index reflecting overall productivity. Foundational contributions by Sweeney and Abadi et al. on K-Anonymity [53] and deep learning with differential privacy [54], respectively, exemplify impactful single papers. Meanwhile, Jin Li stands out for the highest h-index, indicative of substantial impact from well-cited papers across the breadth of this research domain.

5) WHAT RESEARCH GAPS AND CHALLENGES CAN BE IDENTIFIED WITHIN AI PRIVACY PROTECTION RESEARCH? ARE THERE EMERGING SUBFIELDS OR NICHE AREAS?

Drawing from the key findings presented, the authors identify several research gaps and challenges within the realm of privacy protection in AI environments.

a: UNDERDEVELOPED INTERDISCIPLINARY COLLABORATION

Although research in this field spans a wide range, covering machine learning, data analysis, distributed systems, cryptography, network security, privacy-preserving techniques, human factors and compliance, the analysis indicates a pattern of insufficient interconnectivity between each theme. Specifically, keywords related to privacy policy and legal frameworks, such as ''GDPR'' and ''privacy policy'', appear in the highly developed and isolated themes quadrant. This suggests that there is an underdeveloped interdisciplinary collaboration between privacy policies and AI technologies.

b: HUMAN FACTORS AND COMPLIANCE LAG

Privacy protection intricately links with human factors and compliance, especially concerning ethical and legal aspects. However, in comparison to the well-developed research in privacy protection from a technical perspective, studies focusing on human factors and compliance appear fragmented and less cohesive, often confined to a narrow context. Compared to keywords in the above technical areas, keywords in these areas have the lowest average links, lowest average link strength, lowest average occurrence, lowest average citations, and a lower average publication year, suggesting that this is a less mature or less established research area. The research in these areas is more incentivized by specific events, such as the enforcement of GDPR and the privacy concerns raised by the global COVID-19 pandemic. Additionally, this gap may result from a reactive research approach, with ethical and legal investigations often trailing behind rapid advancements in technical solutions.

c: GLOBAL IMBALANCE IN RESEARCH OUTPUT

Whereas the top 10 countries have been prominent in advancing research within the realm of privacy protection in AI, there is a notable gap in contributions from regions like Africa and South America. The research also indicates that geographical proximity influences collaboration patterns worldwide.

6) WHAT RECOMMENDATIONS CAN ENHANCE FUTURE RESEARCH ON PRESERVING PRIVACY IN AI ENVIRONMENTS?

a: INTERDISCIPLINARY COLLABORATION FOR COMPREHENSIVE PRIVACY SOLUTIONS

Privacy protection in AI is inherently multidisciplinary, requiring collaboration among AI researchers, social scientists, psychologists, policy experts, educators, legal professionals, ethicists, and other stakeholders. Encouraging such collaboration can significantly contribute to addressing complex privacy challenges comprehensively. Recognizing the current singular perspectives in AI and privacy protection publications, there is a need for a more holistic approach in the development of solutions. In the era of big data, safeguarding user data privacy requires a collective effort to mitigate risks and create safer AI environments. Understanding the current state of research, identifying key contributors, and recognizing emerging themes are vital for expanding knowledge boundaries.

b: RESEARCH FOCUS ON HUMAN FACTORS AND COMPLIANCE

Emphasizing the ethical and legal dimensions of privacy protection within AI environments is crucial. Research efforts should concentrate on areas such as informed consent, algorithmic bias and fairness, legal frameworks and compliance, transparency and explainability, accountability and liability, and international collaboration on legal standards. These areas play a pivotal role in establishing privacy-protected AI environments, empowering individuals to have greater control over their personal data within AI ecosystems. Addressing these dimensions is essential for the responsible development and deployment of AI systems.

c: PROMOTING INTERNATIONAL COLLABORATION FOR HOLISTIC PRIVACY PROTECTION

In the era of globalized AI products and systems, safeguarding personal privacy from AI risks necessitates an emphasis on international collaboration. A multi-faceted approach, considering cultural, economic, and educational factors, is crucial, particularly in regions with less research activity. Conducting case studies in underrepresented areas can advance the understanding of privacy challenges, and fostering collaboration on a global scale is essential for sharing insights across diverse cultural backgrounds. Bridging geographical research gaps is vital for a systematic understanding of region-specific privacy challenges, and identifying mechanisms to encourage collaboration among a more diverse set of countries is a key research challenge.

B. LIMITATIONS

The analysis relies on data from specific bibliographic databases, which may not comprehensively cover all relevant publications, potentially leading to the omission of significant research contributions from sources not included in the dataset. Research has shown that whereas Web of Science (WoS) and Scopus are widely respected bibliographic databases for academic research, they tend to heavily favor specific scientific disciplines. This bias tends to prioritize fields such as Natural Sciences, Engineering, and Biomedical Research whereas potentially underrepresenting the Social Sciences and Arts and Humanities [95]. The bibliometric analysis discussed earlier solely relies on the Web of Science (WoS) dataset as its primary source of scientific

publications. Consequently, this exclusive reliance could restrict the analysis from offering a systematic view of the entire academic landscape. Furthermore, the analysis is primarily based on publications in English, which could result in the exclusion of valuable research published in other languages. Additionally, whereas efforts were made to use extensive search terms, the choice of keywords and search terms used to retrieve publications may introduce bias, possibly causing relevant papers that employ different terminology to be overlooked.

V. CONCLUSION

This article provides a systematic analysis of privacy protection in AI environments, revealing a dynamic landscape marked by notable contributions, evolving trends, and emerging challenges.

The study underscores the heightened attention to privacy protection in AI, particularly gaining momentum after 2018. Pioneering sources such as *IEEE Transactions on Knowledge and Data Engineering* and influential journals like *IEEE Access* and the *IEEE Internet of Things Journal* have played pivotal roles. Foundational works by Sweeney on K-Anonymity [53] and Abadi et al.'s contributions to deep learning with differential privacy [54] stand out in this field. Other authors, including Jin Li, have also made substantial contributions, as reflected in their high h-indexes.

The landscape of privacy preservation research has evolved from early algorithm-focused approaches, such as k-anonymity, to data-centric solutions in response to the advent of big data and blockchain. Since the mid-2010s, researchers have begun leveraging cloud resources to enhance privacy-preserving computations. Recently, Federated Learning and Differential Privacy have emerged as primary areas of focus, showcasing the field's adaptability in addressing emerging challenges in privacy-preserving AI.

The study also unveils the global research landscape, with the top 10 countries like China, the USA, India, Australia, Germany, the UK, Japan, Korea, Canada, and Spain collectively contributing to 74.8% of articles in this field. Specifically, the top 10 affiliations with the most publications are predominantly located in China, the US, and Australia. Conversely, regions in Africa and South America exhibit comparatively lower scientific production.

Drawing from these key findings, the authors have identified significant gaps and challenges in privacy protection in AI. These encompass a global imbalance in research output, a lag in research regarding the ethical and legal dimensions of privacy protection, and the necessity for enhanced interdisciplinary collaboration to develop comprehensive privacy solutions. Addressing these gaps and challenges is vital to ensure a robust and inclusive approach to privacy protection in AI.

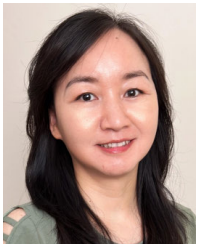
In conclusion, privacy protection in AI is a dynamic and evolving field that necessitates continuous collaboration, interdisciplinary engagement, and a global perspective to comprehensively address emerging challenges. As the world becomes increasingly interconnected, safeguarding user privacy in the era of big data and AI becomes paramount. The authors anticipate that this analysis will contribute to the ongoing discourse and inspire further research and collaboration in the pursuit of a privacy-protected AI environments.

REFERENCES

[1] P. Szolovits, *Artificial Intelligence in Medicine*. Evanston, IL, USA: Routledge, 2019.
[2] L. Cao, "AI in finance: A review," Univ. Technol. Sydney, NSW, Australia, Tech. Rep., Jan. 2020. [Online]. Available: http://dx.doi.org/10.2139/ssrn.3647625
[3] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, "Applications of artificial intelligence in transport: An overview," *Sustainability*, vol. 11, no. 1, p. 189, Jan. 2019.
[4] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artif. Intell. Agricult.*, vol. 2, pp. 1–12, Jun. 2019.
[5] S. Hajkowicz, C. Sanderson, S. Karimi, A. Bratanova, and C. Naughtin, "Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960–2021," *Technol. Soc.*, vol. 74, Aug. 2023, Art. no. 102260.
[6] S. Yu and F. Carroll, "Securing privacy during a world health emergency: Exploring how to create a balance between the need to save the world and people's right to privacy," in *Data Protection in a Post-Pandemic Society: Laws, Regulations, Best Practices and Recent Solutions*. Berlin, Germany: Springer, 2023, pp. 145–167.
[7] D. Zhang, N. Maslej, E. Brynjolfsson, J. Etchemendy, T. Lyons, J. Manyika, H. Ngo, J. C. Niebles, M. Sellitto, E. Sakhaee, Y. Shoham, J. Clark, and R. Perrault, "Artificial intelligence index report 2022," AI Index Steering Committee, Stanford Inst. Hum.-Centered AI, Stanford Univ., California, Tech. Rep., Mar. 2022. [Online]. Available: https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf
[8] K. Manheim and L. Kaplan, "Artificial intelligence: Risks to privacy and democracy," *Yale JL Tech.*, vol. 21, pp. 106–188, Jan. 2019.
[9] C. Wang, Y. Zheng, J. Jiang, and K. Ren, "Toward privacy-preserving personalized recommendation services," *Engineering*, vol. 4, no. 1, pp. 21–28, Feb. 2018.
[10] A. Taeihagh and H. S. M. Lim, "Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks," *Transp. Rev.*, vol. 39, no. 1, pp. 103–128, Jan. 2019.
[11] T. Bolton, T. Dargahi, S. Belguith, M. S. Al-Rakhami, and A. H. Sodhro, "On the security and privacy challenges of virtual assistants," *Sensors*, vol. 21, no. 7, p. 2312, Mar. 2021.
[12] G. A. Kaissis, M. R. Makowski, D. Ruckert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Mach. Intell.*, vol. 2, no. 6, pp. 305–311, Jun. 2020.
[13] UN News. (Nov. 2021). *193 Countries Adopt First-Ever Global Agreement on the Ethics of Artificial Intelligence*. [Online]. Available: https://news.un.org/en/story/2021/11/1106612
[14] United Nations Educational Scientific and Cultural Organization (UNESCO). (2021). *Recommendation on the Ethics of Artificial Intelligence*. Accessed: Dec. 2, 2024. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137/PDF/381137eng.pdf.multi
[15] OECD AI Policy Observatory. (2023). *National AI Policies*. Accessed: Nov. 2, 2024. [Online]. Available: https://oecd.ai/en/wonk/national-policies-2
[16] The White House. (Oct. 2023). *Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence*. Accessed: Nov. 2, 2024. [Online]. Available: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/
[17] European Parliament. (Dec. 2021). *Artificial Intelligence Act*. Accessed: Feb. 14, 2024. [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf
[18] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *J. Bus. Res.*, vol. 133, pp. 285–296, Sep. 2021.

[19] Z. Ünal, "Smart farming becomes even smarter with deep learning—A bibliographical analysis," *IEEE Access*, vol. 8, pp. 105587–105609, 2020.
[20] J.-C. Liang, G.-J. Hwang, M.-R.-A. Chen, and D. Darmawansah, "Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach," *Interact. Learn. Environ.*, vol. 31, no. 7, pp. 4270–4296, Oct. 2023.
[21] C. Ma, Q. Xu, and B. Li, "Comparative study on intelligent education research among countries based on bibliographic coupling analysis," *Library Hi Tech.*, vol. 40, no. 3, pp. 786–804, May 2022.
[22] A. Entezari, A. Aslani, R. Zahedi, and Y. Noorollahi, "Artificial intelligence and machine learning in energy systems: A bibliographic perspective," *Energy Strategy Rev.*, vol. 45, Jan. 2023, Art. no. 101017.
[23] M. M. Islam, T. N. Poly, B. Alsinglawi, L.-F. Lin, S.-C. Chien, J.-C. Liu, and W.-S. Jian, "Application of artificial intelligence in COVID-19 pandemic: Bibliometric analysis," *Healthcare*, vol. 9, no. 4, p. 441, Apr. 2021.
[24] C. M. Feng, A. Park, L. Pitt, J. Kietzmann, and G. Northey, "Artificial intelligence in marketing: A bibliographic perspective," *Australas. Marketing J.*, vol. 29, no. 3, pp. 252–263, Aug. 2021.
[25] Z. H. Munim, M. Dushenko, V. J. Jimenez, M. H. Shakil, and M. Imset, "Big data and artificial intelligence in the maritime industry: A bibliometric review and future research directions," *Maritime Policy Manag.*, vol. 47, no. 5, pp. 577–597, Jul. 2020.
[26] A. K. Shukla, M. Janmaijaya, A. Abraham, and P. K. Muhuri, "Engineering applications of artificial intelligence: A bibliometric analysis of 30 years (1988–2018)," *Eng. Appl. Artif. Intell.*, vol. 85, pp. 517–532, Oct. 2019.
[27] A. Darko, A. P. C. Chan, M. A. Adabre, D. J. Edwards, M. R. Hosseini, and E. E. Ameyaw, "Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities," *Autom. Construct.*, vol. 112, Apr. 2020, Art. no. 103081.
[28] J. W. Goodell, S. Kumar, W. M. Lim, and D. Pattnaik, "Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis," *J. Behav. Experim. Finance*, vol. 32, Dec. 2021, Art. no. 100577.
[29] M. A. Agustí and M. Orta-Pérez, "Big data and artificial intelligence in the fields of accounting and auditing: A bibliometric analysis," *Spanish J. Finance Accounting/Revista Española de Financiación y Contabilidad*, vol. 52, no. 3, pp. 412–438, Jul. 2023, doi: 10.1080/02102412.2022.2099675.
[30] A. S. Ali, Z. F. Zaaba, and M. M. Singh, "Privacy during epidemic of COVID-19: A bibliometric analysis," *Bull. Electr. Eng. Informat.*, vol. 12, no. 1, pp. 587–596, Feb. 2023.
[31] T. Saheb, T. Saheb, and D. O. Carpenter, "Mapping research strands of ethics of artificial intelligence in healthcare: A bibliometric and content analysis," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104660.
[32] Y. Zhang, M. Wu, G. Y. Tian, G. Zhang, and J. Lu, "Ethics and privacy of artificial intelligence: Understandings from bibliometrics," *Knowl.-Based Syst.*, vol. 222, Jun. 2021, Art. no. 106994.
[33] S. Fosso Wamba, R. E. Bawack, C. Guthrie, M. M. Queiroz, and K. D. A. Carillo, "Are we preparing for a good AI society? A bibliometric review and research agenda," *Technol. Forecasting Social Change*, vol. 164, Mar. 2021, Art. no. 120482.
[34] Y. Yin, D. Chun, Z. Tang, and M. Huang, "A comparative analysis of the current status and trends of domestic and international privacy protection research—CiteSpace-based bibliometric study (1976–2022)," *Open J. Bus. Manage.*, vol. 10, no. 6, pp. 3024–3047, 2022.
[35] J. Zhu and W. Liu, "A tale of two databases: The use of web of science and scopus in academic papers," *Scientometrics*, vol. 123, no. 1, pp. 321–335, Apr. 2020.
[36] N. Liu, P. Shapira, and X. Yue, "Tracking developments in artificial intelligence research: Constructing and applying a new search strategy," *Scientometrics*, vol. 126, no. 4, pp. 3153–3192, Apr. 2021.
[37] A. Pritchard, "Statistical bibliography or bibliometrics?" *J. Document.*, vol. 25, p. 348, Jan. 1969.
[38] E. C. M. Noyons, H. F. Moed, and M. Luwel, "Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study," *J. Amer. Soc. Inf. Sci.*, vol. 50, no. 2, pp. 115–131, 1999.
[39] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 46, pp. 16569–16572, Nov. 2005. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0507655102
[40] L. Egghe, "Theory and practise of the g-index," *Scientometrics*, vol. 69, no. 1, pp. 131–152, Oct. 2006.
[41] J. E. Hirsch, "Does the h index have predictive power?" *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 49, pp. 19193–19198, Dec. 2007, doi: 10.1073/pnas.0707962104.
[42] W. M. Sweileh, A. S. AbuTaha, A. F. Sawalha, S. Al-Khalil, S. W. Al-Jabi, and S. H. Zyoud, "Bibliometric analysis of worldwide publications on multi-, extensively, and totally drug—Resistant tuberculosis (2006–2015)," *Multidisciplinary Respiratory Med.*, vol. 11, pp. 1–16, Jan. 2017.
[43] H. Small, "Update on science mapping: Creating large document spaces," *Scientometrics*, vol. 38, no. 2, pp. 275–293, Feb. 1997.
[44] S. A. Morris and B. Van der Veer Martens, "Mapping research specialties," *Annu. Rev. Inf. Sci. Technol.*, vol. 42, no. 1, pp. 213–295, Jan. 2008.
[45] H. P. F. Peters and A. F. J. Van Raan, "Structuring scientific activities by co-author analysis: An expercise on a university faculty level," *Scientometrics*, vol. 20, no. 1, pp. 235–255, Jan. 1991.
[46] H. Small, "Co-citation in the scientific literature: A new measure of the relationship between two documents," *J. Amer. Soc. Inf. Sci.*, vol. 24, no. 4, pp. 265–269, Jul. 1973.
[47] E. Garfield, "Historiographic mapping of knowledge domains literature," *J. Inf. Sci.*, vol. 30, no. 2, pp. 119–145, Apr. 2004.
[48] M. Aria, C. Cuccurullo, L. D'Aniello, M. Misuraca, and M. Spano, "Thematic analysis as a new culturomic tool: The social media coverage on COVID-19 pandemic in Italy," *Sustainability*, vol. 14, no. 6, p. 3643, Mar. 2022.
[49] M. Aria, M. Misuraca, and M. Spano, "Mapping the evolution of social research and data science on 30 years of social indicators research," *Social Indicators Res.*, vol. 149, no. 3, pp. 803–831, Jun. 2020.
[50] A. Perianes-Rodriguez, L. Waltman, and N. J. van Eck, "Constructing bibliometric networks: A comparison between full and fractional counting," *J. Informetrics*, vol. 10, no. 4, pp. 1178–1195, Nov. 2016.
[51] M. Aria and C. Cuccurullo, "Bibliometrix : An R-tool for comprehensive science mapping analysis," *J. Informetrics*, vol. 11, no. 4, pp. 959–975, Nov. 2017.
[52] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010.
[53] L. Sweeney, "K-anonymity: A model for protecting privacy," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 10, no. 5, pp. 557–570, Oct. 2002, doi: 10.1142/s0218488502001648.
[54] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.* New York, NY, USA: Association for Computing Machinery, Oct. 2016, pp. 308–318, doi: 10.1145/2976749.2978318.
[55] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.* New York, NY, USA: Association for Computing Machinery, Oct. 2017, pp. 1175–1191, doi: 10.1145/3133956.3133982.
[56] P. Mohassel and Y. Zhang, "SecureML: A system for scalable privacy-preserving machine learning," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2017, pp. 19–38.
[57] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*. New York, NY, USA: Association for Computing Machinery, Sep. 2015, pp. 909–910, doi: 10.1109/ALLERTON.2015.7447103.
[58] L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai, "Privacy-preserving deep learning via additively homomorphic encryption," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 5, pp. 1333–1345, May 2018.
[59] A. B. Chan, Z.-S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–7, doi: 10.1109/CVPR.2008.4587569.
[60] G. Bansal and D. Gefen, "The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online," *Decis. Support Syst.*, vol. 49, no. 2, pp. 138–150, May 2010.
[61] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. S. Quek, and H. V. Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3454–3469, 2020.
[62] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4177–4186, Jun. 2020.

[63] G. Melin and O. Persson, "Studying research collaboration using co-authorships," *Scientometrics*, vol. 36, no. 3, pp. 363–377, Jul. 1996. [Online]. Available: https://api.semanticscholar.org/CorpusID:20236217
[64] M. Kantarcioglu and C. Clifton, "Privacy-preserving distributed mining of association rules on horizontally partitioned data," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 9, pp. 1026–1037, Sep. 2004.
[65] E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 2, pp. 232–243, Feb. 2005.
[66] K. Liu, H. Kargupta, and J. Ryan, "Random projection-based multiplicative data perturbation for privacy preserving distributed data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 1, pp. 92–106, Jan. 2006.
[67] J. Vaidya, M. Kantarcioglu, and C. Clifton, "Privacy-preserving naive Bayes classification," *VLDB J.*, vol. 17, no. 4, pp. 879–898, 2008.
[68] V. Nikolaenko, U. Weinsberg, S. Ioannidis, M. Joye, D. Boneh, and N. Taft, "Privacy-preserving ridge regression on hundreds of millions of records," in *Proc. IEEE Symp. Secur. Privacy*, May 2013, pp. 334–348.
[69] J. Yuan and S. Yu, "Privacy preserving back-propagation neural network learning made practical with cloud computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 212–221, Jan. 2014.
[70] Q. Zhang, L. T. Yang, and Z. Chen, "Privacy preserving deep computation model on cloud for big data feature learning," *IEEE Trans. Comput.*, vol. 65, no. 5, pp. 1351–1362, May 2016.
[71] P. Li, J. Li, Z. Huang, T. Li, C.-Z. Gao, S.-M. Yiu, and K. Chen, "Multi-key privacy-preserving deep learning in cloud computing," *Future Gener. Comput. Syst.*, vol. 74, pp. 76–85, Sep. 2017.
[72] N. Phan, X. Wu, H. Hu, and D. Dou, "Adaptive Laplace mechanism: Differential privacy preservation in deep learning," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2017, pp. 385–394.
[73] M. Nasr, R. Shokri, and A. Houmansadr, "Machine learning with membership privacy using adversarial regularization," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.* New York, NY, USA: Association for Computing Machinery, Oct. 2018, pp. 634–646, doi: 10.1145/3243734.3243855.
[74] M. Al-Rubaie and J. M. Chang, "Privacy-preserving machine learning: Threats and solutions," *IEEE Secur. Privacy*, vol. 17, no. 2, pp. 49–58, Mar. 2019.
[75] N. Wang, X. Xiao, Y. Yang, J. Zhao, S. C. Hui, H. Shin, J. Shin, and G. Yu, "Collecting and analyzing multidimensional data with local differential privacy," in *Proc. IEEE 35th Int. Conf. Data Eng. (ICDE)*, Apr. 2019, pp. 638–649.
[76] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2019, pp. 739–753.
[77] P. C. Mahawaga Arachchige, P. Bertok, I. Khalil, D. Liu, S. Camtepe, and M. Atiquzzaman, "Local differential privacy for deep learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5827–5842, Jul. 2020.
[78] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, "Efficient and privacy-enhanced federated learning for industrial artificial intelligence," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6532–6542, Oct. 2020.
[79] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Gener. Comput. Syst.*, vol. 115, pp. 619–640, Feb. 2021.
[80] Q. He, "Knowledge discovery through co-word analysis," *Library Trends*, vol. 48, no. 1, pp. 59–133, 1999.
[81] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955," *AI Mag.*, vol. 27, no. 4, p. 12, 2006.
[82] M. Haenlein and A. Kaplan, "A brief history of artificial intelligence: On the past, present, and future of artificial intelligence," *California Manag. Rev.*, vol. 61, no. 4, pp. 5–14, Aug. 2019.
[83] S. Naker and D. Greenbaum, "Now you see me: Now you still do: Facial recognition technology and the growing lack of privacy," *BUJ Sci. Tech. L.*, vol. 23, p. 88, Jan. 2017.
[84] N. J. Fast and A. S. Jago, "Privacy matters... or does It? Algorithms, rationalization, and the erosion of concern for privacy," *Current Opinion Psychol.*, vol. 31, pp. 44–48, Feb. 2020.
[85] E. Alepis and C. Patsakis, "Monkey says, monkey does: Security and privacy on voice assistants," *IEEE Access*, vol. 5, pp. 17841–17851, 2017.
[86] H. Lin and N. Bergmann, "IoT privacy and security challenges for smart home environments," *Information*, vol. 7, no. 3, p. 44, Jul. 2016.
[87] B. Murdoch, "Privacy and artificial intelligence: Challenges for protecting health information in a new era," *BMC Med. Ethics*, vol. 22, no. 1, pp. 1–5, Dec. 2021.
[88] J. Hinds, E. J. Williams, and A. N. Joinson, "'It wouldn't happen to me': Privacy concerns and perspectives following the Cambridge Analytica scandal," *Int. J. Hum.-Comput. Stud.*, vol. 143, Nov. 2020, Art. no. 102498.
[89] I. N. Rezende, "Facial recognition in police hands: Assessing the 'clearview case' from a European perspective," *New J. Eur. Criminal Law*, vol. 11, no. 3, pp. 375–389, Sep. 2020.
[90] A. Pfeifle, "Alexa, what should we do about privacy: Protecting privacy for users of voice-activated devices," *Wash. L. Rev.*, vol. 93, pp. 421–458, Jan. 2018.
[91] S. Sharma, *Data Privacy and GDPR Handbook*. Hoboken, NJ, USA: Wiley, 2019.
[92] T. Linden, R. Khandelwal, H. Harkous, and K. Fawaz, "The privacy policy landscape after the GDPR," 2019, *arXiv:1809.08396*.
[93] S. Yu and F. Carroll, "Implications of AI in national security: Understanding the security issues and ethical challenges," in *Artificial Intelligence in Cyber Security: Impact and Implications: Security Challenges, Technical and Ethical Issues, Forensic Investigative Challenges*. Berlin, Germany: Springer, 2022, pp. 157–175.
[94] *Artificial Intelligence*. Accessed: Mar. 15, 2024. [Online]. Available: https://new.nsf.gov/focus-areas/artificial-intelligence
[95] P. Mongeon and A. Paul-Hus, "The journal coverage of web of science and scopus: A comparative analysis," *Scientometrics*, vol. 106, no. 1, pp. 213–228, Jan. 2016.



SHASHA YU received the J.M. degree from Xiangtan University, China, in 2006, the M.Sc. degree in advanced computer science from Cardiff Metropolitan University, U.K., in 2022, and the M.S. degree in data analytics from Clark University, USA, in 2023. Her research interests include artificial intelligence, digital policing, privacy protection, information security, and cybersecurity.

FIONA CARROLL (Senior Member, IEEE) received the Ph.D. degree in computing from Edinburgh Napier University, Scotland, in 2008. She is currently a Reader with the School of Technologies, Cardiff Metropolitan University, Wales. Her research over the past 17 years has focused on the fast changing relations between humans and digital technologies. It is multi-disciplinary and shows a substantial contribution to scholarship in the fields of human computer interaction. In particular, her interests lie in aesthetic and ethical computing, human dimensions of cyber security, and educational technologies. Her current research explores the aesthetic of cyberspace and how we might design for the immediate informed "sense" that people need to prevent and/or cope with harm online. As an accomplished academic, she has numerous peer reviewed publications and she has partaken in cutting-edge research projects both at national and international levels.

BARRY L. BENTLEY received the Ph.D. degree in biological science from the University of Cambridge, U.K., the P.G.Cert. degree in history from the University of Oxford, U.K., and the B.Sc. degree in computer science from the University of Plymouth, U.K. He is currently a Reader in bio-engineering with Cardiff Metropolitan University, U.K., having been a member of the Cardiff School of Technologies, since 2021. He is also a 2023–2024 Fulbright Scholar with the Harvard Medical School, USA, and a fellow of UCL's Collaboration for the Advancement of Sustainable Medical Innovation (CASMI), U.K., and the Foresight Institute, San Francisco, USA. His primary research interests include the development of technologies to accelerate biomedical research, including the development of classification and staging systems for age-related pathology.