

A Comparative Analysis of the Current Status and Trends of Domestic and International Privacy Protection Research—CiteSpace-Based Bibliometric Study (1976-2022)

Yushi Yin^{1,2}, Dongphil Chun¹, Zhiying Tang², Minghao Huang^{3*}

¹Graduate School of Management of Technology, Pukyong National University, Busan, South Korea

²College of Tropical Crops, Yunnan Agricultural University, Puer, China

³Seoul School of Integrated Science & Technologies, Seoul, South Korea

Email: yinyushi7@gmail.com, *mhuang@assist.ac.kr

How to cite this paper: Yin, Y. S., Chun, D., Tang, Z. Y., & Huang, M. H. (2022). A Comparative Analysis of the Current Status and Trends of Domestic and International Privacy Protection Research—CiteSpace-Based Bibliometric Study (1976-2022). *Open Journal of Business and Management*, 10, 3024-3047.

<https://doi.org/10.4236/ojbm.2022.106150>

Received: October 19, 2022

Accepted: November 5, 2022

Published: November 8, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Objective: This paper analyzes the literature on the topic of privacy protection based on the core databases included in China Knowledge Network (CNKI) and Web of Science (WOS) from 1976 to 2022, and provides reference for the subsequent research on privacy protection in China by comparing and summarizing the status of domestic and foreign research. **Methods:** The Web of science core collection database was used as the foreign data source, and the China Knowledge Network database was used as the domestic data source to retrieve the relevant literature built up to August 2, 2022, and the authors, countries, institutions, and keywords were compared and visualized by CiteSpace 6.1.R2 (64-bit). **Results:** A total of 8223 English-language and 4573 Chinese-language papers were included, and the number of relevant studies both at home and abroad was on the rise. Foreign scholars and institutions cooperate closely with each other and build a stable cooperation network; while the number of domestic research teams is small and cooperation is limited. Privacy concern and privacy protection are common research focuses at home and abroad, while foreign research focuses on online privacy concern and domestic research pays more attention to computer science privacy protection model. Privacy information is a current international research hotspot, and domestic AI privacy protection concerns are high. **Conclusion:** Domestic and foreign research in this field is increasingly concerned but each has its own focus, and domestic research has problems such as lack of cooperation and lagging research that need to be improved, and privacy protection is a direction worthy of in-depth research in the future.

Keywords

Privacy Protection, Comparative Analysis, Bibliography, CiteSpace

1. Introduction

Digital marketplaces have been understood as social networks, large marketplaces, and any digital platform that identifies individual user traffic in the form of online communities (Liu & Lai, 2022; Oberg & Alexander, 2019). If these datasets are studied using artificial intelligence, machine learning, deep learning, or big data (BDA) analysis (Kar & Dwivedi, 2020), user behavior can be better predicted, but also increases the risk of user privacy violations in the digital ecosystem (Gutierrez et al., 2019). Artificial intelligence fundamentally changes the way human production and life, and a freer, more open and transparent space is revealed (Xu, Xiang, & He, 2021). With leaps and bounds in information technology, the epidemic has enabled people to work, socialize, and shop online, but inadvertently increased privacy concerns and personal privacy protection behaviors (Sundaram & Shetty, 2022). Privacy has never been more important or eroded than now, and information privacy security has become the most important ethical, legal, social, and political issue of our time (Ozdemir, Jeff Smith, & Benamati, 2017), and privacy protection needs urgent attention and improvement.

Privacy protection has been studied by scholars using CiteSpace, and the earliest documented in China is (Zhu, 2013), revealing its development patterns and trends, limited to the extremely small amount of privacy protection literature at that time, using only foreign data for analysis. Later (Gu & Fan, 2018; Feng, Yan, & Hu, 2020; Wang, 2021), etc. only based on domestic or (Xu, Hou, & Hu, 2020; Shu & Liu, 2021) only based on foreign data for research in the field of privacy, the research topic focuses on privacy, which is broader and lacks comparative research. Therefore, based on the unfocused content and incomplete analysis of previous studies, this study uses CiteSpace to analyze the existing literature related to privacy protection at home and abroad, reveal the dynamics of privacy protection research, and provide a thinking path for researchers to understand the current state of research and carry out in-depth research (Table 1).

CiteSpace (Citation Space), developed by the internationally renowned information visualization expert Professor Chaomei Chen, analyzes the potential knowledge contained in the vast scientific literature in the context of scientometrics, data and information visualization, with multiple, time-sharing and dynamic citation visualization. Through a clever spatial layout, the evolution of the field is centrally presented on the knowledge graph of the citation network, and the research frontiers characterized by the citation node literature and co-citation

clustering on the graph as the basis of knowledge are automatically identified, and the graph is highly interpretable (Li & Chen, 2017).

Table 1. Privacy-preserving research articles based on the CiteSpace Bibliometric approach.

Serial number	Year	Article Level	Author	Data source	Data volume/article	Title	Analysis content
1	2013	CNKI Core	Zhu Wei	SCI, SSCI	174	A survey and analysis of domestic and international research on social network privacy protection	Era, country, region, author, issuing institution, discipline of the document
2	2018	CNKI Core	Gu Liping Fan Haichao	CNKI Core	719	The Academic Field of Ten Years of Research on Online Privacy	Number of annual publications, number of articles in grant programs, disciplinary distribution, author co-occurrence, keyword co-occurrence, keyword clustering, hot time zones
3	2020	CNKI Core	Feng Yafei, Yan Chun, Hu Changping	CNKI Core	3861	Structural features and characteristics of domestic privacy research in the past 20 years	Year, discipline, author co-authorship, co-word, emergent word, hot time zone map, keyword clustering
4	2020	CNKI Core	Xu Jinghong, Hou Weipeng, Hu Shiming	SCI, SSCI, A & HCI	13361	Historical Changes and Frontiers of Privacy Research in English Language Academia over the Past Century	Research themes, keywords, critical literature, emergent literature, disciplinary distribution
5	2021	CNKI Core	Wang Hao	CNKI	115	Research on the Rise of Personal Privacy Protection Awareness in the Era of Big Data	Volume of articles, institutional cooperation, keywords
6	2021	SCI	Sheng Shu, and Yi Liu	SCI, SSCI, A & HCI	2532	Looking Back to Move Forward: A Bibliometric Analysis of Consumer Privacy Research	Region, institution, author, number of publications, emergent words

2. Information and Methods

2.1. Data Sources

In China, the data sources were “Core Journals”, “CSSCI” and “CSCD”, and the relevant literature from the beginning of the database to August 2, 2022 was searched with the following search formula The search formula was “Subject = privacy protection or privacy concern”, and the language was limited to “Chinese” (Figure 1). The data of Web of science core collection database science citation index (SCI), social sciences citation index (SSCI) were used abroad, and the relevant literature from the beginning of database construction to August 2, 2022 was retrieved in order to study the privacy protection issue more comprehensively. The search formula of Web of science is TI = “privacy protect*” OR TI = “privacy concern*”, the language was limited to “English”, the type of article was “Article” AND “Review Article The language is limited to “English” and the type of article is “Article” and “Review Article” (Figure 2).



Figure 1. Chinese literature keyword search result chart.

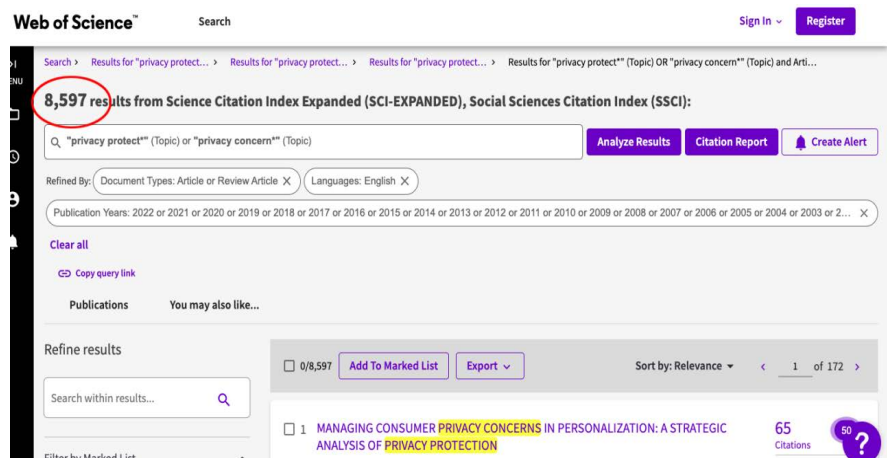


Figure 2. English literature keyword search result chart.

2.2. Adoption and Exclusion Criteria

The criteria for adoption were the relevance of privacy protection and privacy concern studies, while those with missing key information, those judged irrelevant by multiple researchers after reading and discussing, and those with conference abstracts, press releases, litigation documents, editorial materials, etc. were excluded and not adopted.

2.3. Data Processing Analysis

The plain text format (.txt) of the English and Chinese literature was exported and saved, and the researchers removed them according to the adoption and exclusion criteria, after which CiteSpace 6.1.R2 (64-bit) software was used to transform and remove duplicate literature (Figure 3, Figure 4), analyze the number

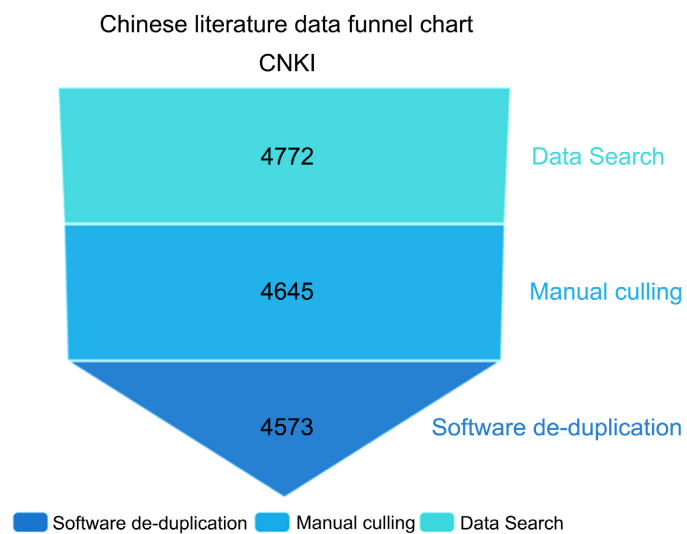


Figure 3. Chinese literature data funnel diagram.

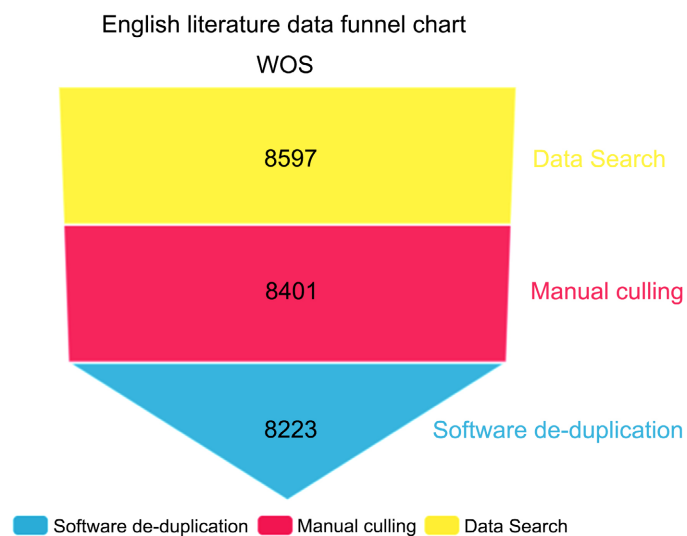


Figure 4. English literature data funnel diagram.

of publications, keywords, national collaboration networks, journal co-citations, etc., and draw visual maps. Parameter setting index: According to the privacy protection and privacy concern themes of Chinese and English literature to build the library to start the search, the time range of English literature was set to 1976-2022, and Chinese to 1997-2022. The time period separation was set to 1 year, and the data inclusion criterion was “TOP N = 50” (the data with the top 50 occurrences in each time slice). In order to avoid the complicated content of the plots, “pruning sliced networks”, “pruning the merged network” and “path-finder” were checked to simplify the plots, and other parameters were left as default values.

2.4. Graph and Metrics Analysis

The node size represents the content of the co-occurrence map. The node size of author collaboration mapping represents the number of papers published by authors, institutions or countries/regions, the node size of network mapping of topics, keywords and scientific categories represents the frequency of occurrence, the size of nodes in co-citation analysis mapping reflects the number of co-citations of literature, authors and journals, and the connecting line reflects the corresponding strength (Li & Chen, 2017). The centrality reflects the closeness of the connection between nodes, and the higher centrality indicates the greater role played by the node.

3. Results

3.1. Annual Publication Volume

The number of publications can measure the hotness of the research field and the progress of the research, and the comparison of the number of publications can more intuitively see the degree of attention and changes in the trend of the research hotspots of the field by researchers at home and abroad. After screening and de-weighting, 4573 and 8223 Chinese and English literature were included respectively, and the annual publication volume is shown in **Figure 5**, with a fit of 0.8844 and 0.974, respectively, and the closer the fit to 1, the higher the credibility of the trend line (Guo & He, 2019). The English literature was first included in 1976, 21 years before the Chinese literature. Both domestic and foreign scholars privacy protection-related literature is growing (due to the inclusion of less literature in 2022, so the number of articles issued in that year decreases), and in recent years there has been an explosive growth abroad, reflecting from the side that scholars are more enthusiastic about the research on privacy protection and privacy concerns.

According to each key point of research in the development of privacy protection, it can be divided into three stages.

1) Initial stage (1976-2004): A small number of scholars at home and abroad researched privacy protection issues in this stage, with an earlier start abroad.

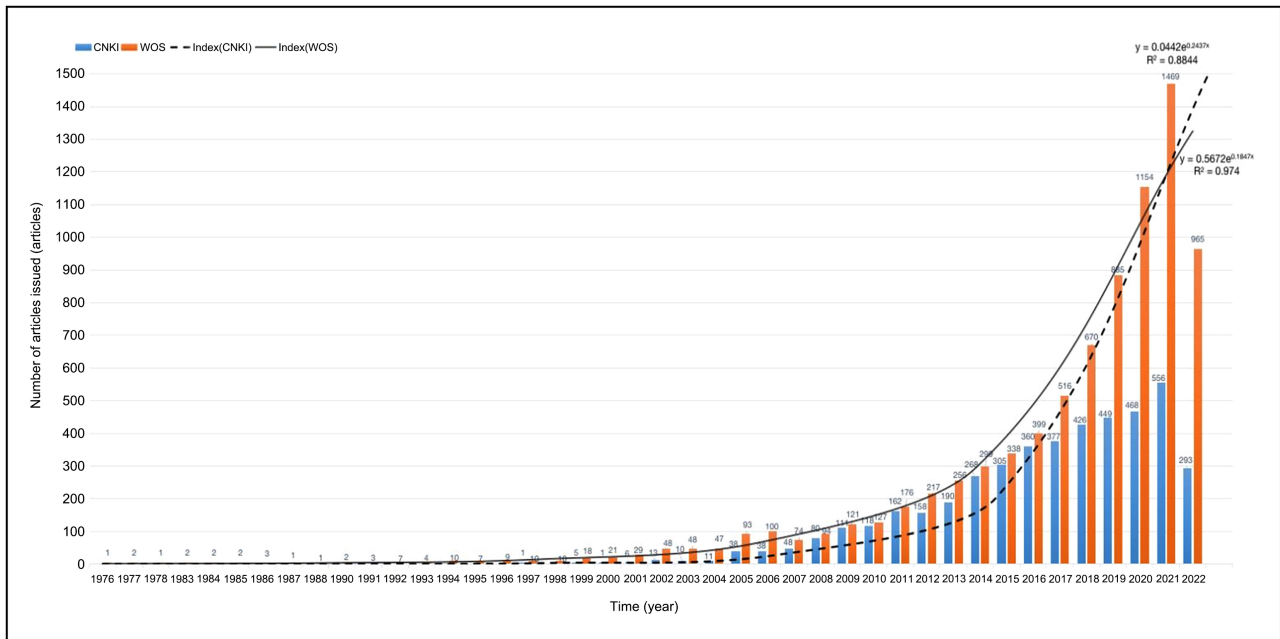


Figure 5. Comparison of the annual publication volume of Chinese and English documents from 1976 to 2022.

The emergence of the Internet had a certain impact on privacy protection, but the relevant knowledge and theories have not yet formed a systematic framework and structure, and WOS and CNKI core collection databases contain less literature, and the annual statistics at home and abroad do not exceed 100 articles. (Minsky, 1976) was the first to conduct privacy protection research, the study pointed out that privacy protection was completely ignored in the database literature, and the programmers maintaining the information system had almost unrestricted access to all the information in it, so deliberate solutions were introduced as an independent dimension of privacy protection.

2) Development phase (2005-2016): this phase is in a stable growth trend at home and abroad, and foreign literature is still more than domestic literature, but the gap between them is not large. The research fields of privacy protection focus on computer science, engineering, telecommunications, and business economics, etc. There are relevant studies on user online privacy, youth privacy issues, and privacy protection models in this phase. The rapid development of information technology and the popular use of e-commerce have led companies to provide personalized services to consumers based on online data. However, the extensive collection and use of personal information has led to serious consumer concerns about privacy invasion and raised the issue of personalized privacy trade-offs (Lee, Ahn, & Bang, 2011).

3) Outbreak phase (2017-2022): The volume of privacy protection publications began to increase dramatically, with foreign countries pulling away from domestic countries by a large margin, and the volume of foreign publications is almost twice as large as domestic ones so far in 2019. 2019 saw the outbreak of the new crown epidemic, with various privacy security issues exposed and the

Covid-19 pandemic increasing privacy protection concerns (Sundaram & Shetty, 2022). China has also accelerated the process of privacy protection-related legislation, continuously introduced relevant national standards, industry standards, and other related management methods, and strengthened the security enforcement inspection of APPs. From the side, it also reflects that scholars' concern for privacy protection has increased substantially in the context of the epidemic.

3.2. Author Analysis

The visual analysis of posting authors can find the core authors as well as collaborating teams in the field of privacy protection research. The circle represents the node, the higher the intermediary centrality of the node the darker the color of the outermost circle will be, the higher the intermediary centrality indicates the greater the role played in the linkage network. Connected lines represent collaborative relationships, and the color and thickness are the time and number of collaborations; the lighter the color, the more recent the time (Rong et al., 2022).

The co-occurrence mapping of authors in Chinese literature is shown in **Figure 6**, which has more sub-networks with more Chinese research groups, but the core teams do not form network linkages with each other and do not cooperate closely. The core team represented by Jing Yang, Bing Zhang, Jing Xie, and Jianpei Zhang, with a surge of 11 publications in 2014, proposed a privacy-preserving algorithm for personalized trajectories based on typical correlation analysis to address the problem of personalized trajectory privacy protection needs, which better respects the privacy protection wishes of data generators and also obtains higher quality trajectories (Li et al., 2015). The core team of Lei Zhang, Chunguang Ma, Bin Wang, Songtao Yang, and Changli Zhou, with a large volume of publications in FY17, studied the privacy-preserving model design established to produce an indistinguishable privacy-preserving method based on the association probability of location offset (Zhang et al., 2017). And the team formed by Yingjie Wu, Weiwei Ni, Geng Chen, and Zhihong Chong analyzed the relationship between preserving the granularity of clustering features and the usability of clustering and the security of privacy protection from the perspective of preserving the granularity of clustering features, and pointed out some difficult problems to be solved and future development directions in the field of clustering-oriented data hiding publishing based on the in-depth comparative analysis of existing technical methods (Ni et al., 2012). The top ten rankings of domestic authors' publications are shown in **Table 2**, with all the centrality being 0. Our authors do not cooperate closely in this field of research. The author with the most publications is Yang Jing, who proposed a privacy-preserving method based on random projection technology to solve the dimensional disaster problem in privacy-preserving data mining (Yang, Zhao, & Zhang, 2013).

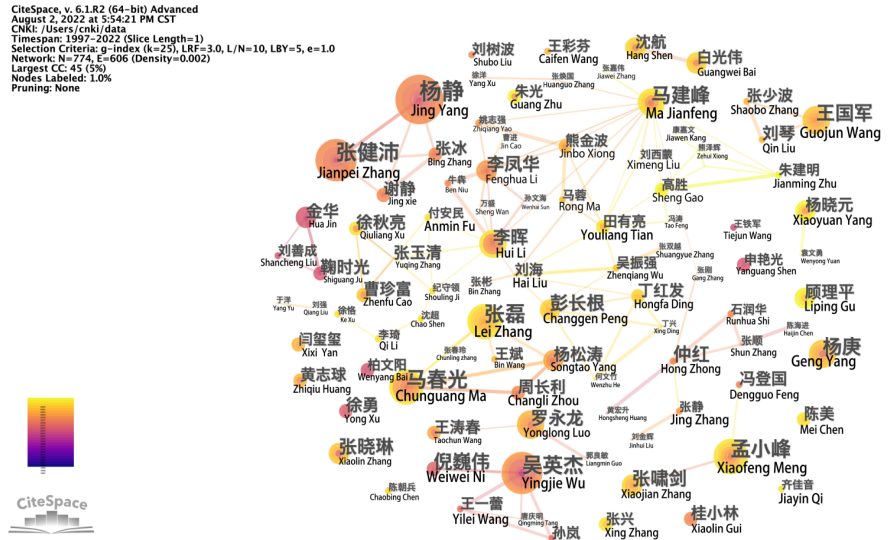


Figure 6. Author’s cooperation network of Chinese literature.

Table 2. Top10 authors of Chinese literature by publishing volume.

Ranking	Author	Number of articles/article	Centrality
1	Jing Yang	32	0.00
2	Jianpei Zhang	27	0.00
3	Lei Zhang	24	0.00
4	Yingjie Wu	23	0.00
5	Chunguang Ma	23	0.00
6	Xiaofeng Meng	21	0.00
7	Geng Yang	20	0.00
8	Jianfeng Ma	19	0.00
9	Weiwei Ni	18	0.00
10	Guojun Wang	17	0.00

The co-occurrence mapping of English literature authors is shown in Figure 7, which shows that the foreign literature authors cooperate very closely and the core research teams are also closely connected with each other. According to the size of the cooperation network and the order of closeness, the volume of publications of English literature authors is nearly 6 times more than that of Chinese literature authors. From Table 3, we can see that the author with the highest total number of publications is ZHANG Y with centrality 0.02. This author’s team

CiteSpace v. 5.1.R2 (64-bit) Advanced
 August 4, 2022 at 5:52:18 PM CST
 WoS: /Users/wos/data
 Timespan: 1976-2022 (Slice Length=1)
 Selection Criteria: g-index (k=5), LRF=3.0, L/N=10, LBY=5, e=1.0
 Network: N=322, E=250 (Density=0.0048)
 Largest CC: 140 (43%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder

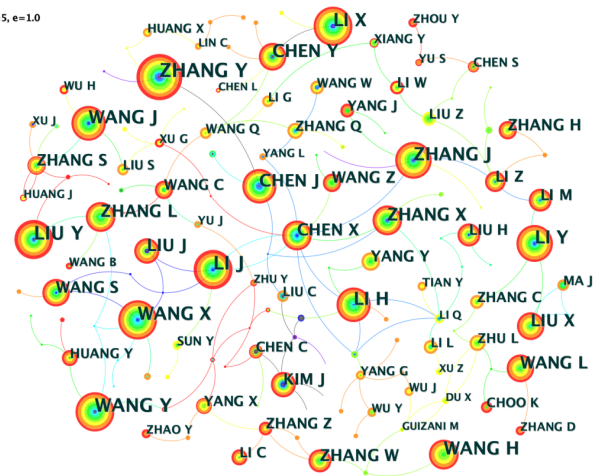


Figure 7. Author's cooperation network of English literature.

Table 3. Top10 authors of English literature by publishing volume.

Ranking	Author	Number of articles/article	Centrality
1	ZHANG Y	166	0.02
2	WANG Y	144	0.02
3	LI J	136	0.03
4	LIU Y	131	0.01
5	WANG X	123	0.02
6	LI Y	123	0.01
7	LI X	120	0.02
8	ZHANG J	119	0.01
9	WANG H	118	0.01
10	LI H	118	0.01

researches ways to protect privacy while providing accountability, and develops the Accountable Monroe Coin model, which isolates the three roles of users, trusted registrars, and trusted regulators. Accountability enables the trusted regulator to reveal the identity of the signer, and only the trusted regulator can track the user's public key as needed (Zhang & Xu, 2022). Notably LI Y saw a spike in the number of articles in 2021, reaching a maximum of 40 articles. The team concentrates on keyword research in privacy protection, and to avoid keyword privacy leakage in both traditional models, the team constructs new global

indexes to protect keyword privacy from insider attackers (Li et al., 2021). Through visual mapping analysis, the study found that the top ten authors of English-language literature are Chinese, mostly collaborating with foreign authors on their publications. In January 2018, the Blue Book on the Development of Chinese Scientific and Technological Journals (2017) counted the inflow and outflow of SCI papers from 14 major paper-producing countries in the world, and China was one of the countries with the most serious loss problem, second only to Portugal. The number of papers published in domestic SCI journals is only 0.12 times the number of papers published by Chinese scholars, meaning that at least 88% of China's SCI articles go to foreign journals (China Association for Science and Technology (ed.), 2018).

3.3. Analysis of Country and Institutional Cooperation Networks

3.3.1. National Cooperation Network

The results were analyzed by combining the names of the countries involved in the web of science core collection source data and their abbreviations, and the results are shown in **Table 4**. The top 5 countries in terms of number of publications are China (3016), the United States (2640), Australia (513), the United Kingdom (464), and South Korea (452), and the highest centrality is the United States (0.34), and the high centrality represents the node then links the size of the role of the entire network. The presence of purple color in the nodes in the map indicates that the country or institution represented has a high intermediary centrality (**Figure 8**).

Table 4. The top ten countries in the national cooperation network.

Ranking	Country/Region	Number of literature/article	Centrality
1	PEOPLES R CHINA	3106	0.05
2	USA	2640	0.34
3	AUSTRALIA	513	0.05
4	ENGLAND	464	0.26
5	SOUTH KOREA	452	0.09
6	CANADA	417	0.04
7	GERMANY	335	0.03
8	TAIWAN CHINA	327	0.03
9	INDIA	281	0.11
10	SINGAPORE	212	0.01

CiteSpace, v. 6.1.R2 (64-bit) Advanced
 August 9, 2022 at 4:47:17 PM CST
 WoS: /Users/wos/data
 Timespan: 2012-2022 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=3.0, L/N=10, LBY=5, e=1.0
 Network: N=108, E=130 (Density=0.0225)
 Largest CC: 104 (96%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder

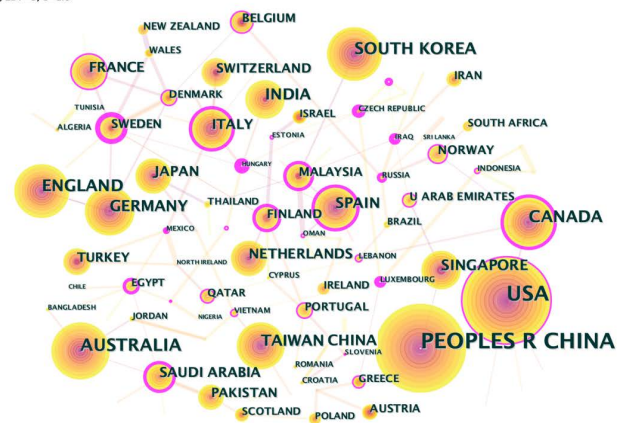


Figure 8. National cooperation network.

3.3.2. Collaboration Network of Issuing Institutions

Collaboration mapping can evaluate the academic influence of countries, institutions or researchers from a new perspective, discover the social relationships between them, and assist us in finding countries, institutions or researchers worthy of attention (Chen et al., 2015). Research from the perspective of regions and overall institutions can better evaluate the research performance and objectively assess the research strengths to complement each other. From **Figure 9**, it can be seen that there is a relatively close cooperation between various research institutions in foreign literature, and only one of the top ten institutions in terms of the number of articles issued is from Singapore (**Table 5**), while the others are from China, represented by Xidian Univ (179 articles issued) and Chinese Acad Sci (centrality 0.05), Xidian University of Electronic Science and Technology is a university with mainly information and electronics disciplines, coordinated development of multiple disciplines, and the first batch of first-class cyber security college construction demonstration projects.

The research institutions of privacy protection Chinese literature (**Figure 10**) also have some cooperative networks, but they are sparse compared to those of English literature, and there is a lack of cooperation among institutions with high centrality (**Table 6**). The research institution with the highest number of publications in Chinese literature is the School of Computer Science and Technology of Nanjing University of Posts and Telecommunications, followed by the School of Computer Science and Technology of Harbin Engineering University, the School of Information Management of Wuhan University, and the School of Information of Renmin University of China, etc. The highest centrality is the Institute of Information Engineering of the Chinese Academy of Sciences, which plays a high role as a bridge. Based on the reading of institutional literature, it is known that the research at the School of Computer Science, Nanjing University of Posts and Telecommunications focuses on the differential privacy protection

mechanism in the context of the Internet of Things and blockchain privacy leakage. With the development of social networks, the security protection of personal privacy data such as users' social relationships needs to be addressed, and the institution has proposed various differential privacy protection schemes, such as BCPA, horizontal federal PCA differential privacy data publishing algorithm, and non-interactive differential privacy model dp-noisy (Huang et al., 2019; Zhu & Yang, 2022; Huang et al., 2020).

Table 5. Top 10 publishing institutions for English literature.

Ranking	Institution	Number of literature/article	Centrality
1	Xidian Univ	179	0.03
2	Chinese Acad Sci	136	0.05
3	Beijing Univ Posts & telecommun	131	0.02
4	Tsinghua Univ	107	0.04
5	Univ Elect Sci & Technol China	99	0.03
6	Wuhan Univ	97	0.03
7	Nanyang Technol Univ	94	0.05
8	Shanghai Jiao Tong Univ	83	0.03
9	Huazhong Univ sci & Technol	77	0.02
10	Zhejiang Univ	77	0.01

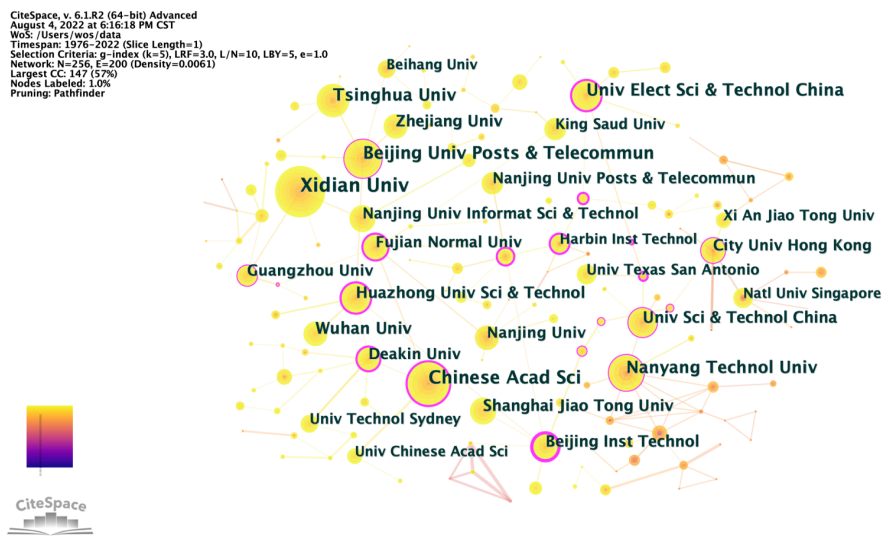


Figure 9. The cooperation network of English-language document publishing organizations.

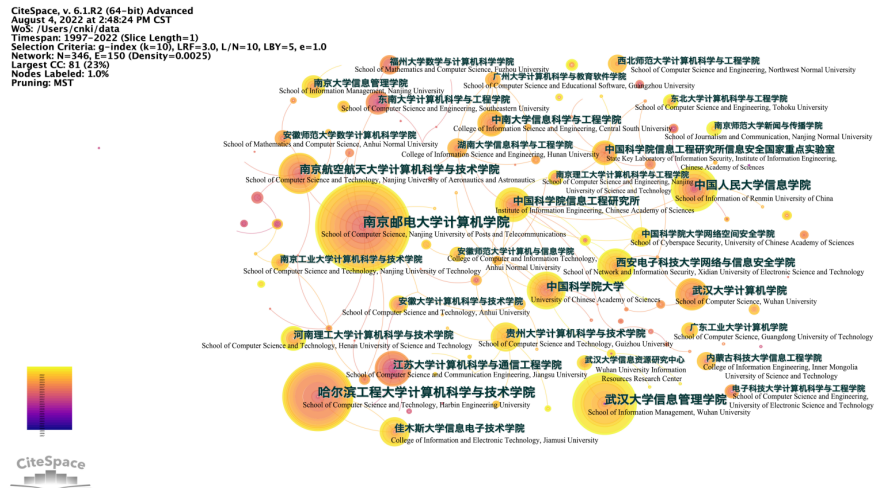


Figure 10. Cooperation network of Chinese document publishing organizations.

Table 6. Top 10 publishing institutions for Chinese literature.

Ranking	Institution	Number of literature/article	Centrality
1	School of Computer Science, Nanjing University of Posts and Telecommunications	86	0.06
2	School of Computer Science and Technology, Harbin Engineering University	64	0.01
3	School of Information Management, Wuhan University	63	0.00
4	School of Information of Renmin University of China	46	0.02
5	School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics	38	0.01
6	School of Computer Science and Communication Engineering, Jiangsu University	38	0.01
7	University of Chinese Academy of Sciences	35	0.02
8	Institute of Information Engineering, Chinese Academy of Sciences	34	0.07
9	School of Computer Science, Wuhan University	33	0.02
10	School of Network and Information Security, Xidian University of Electronic Science and Technology	33	0.02

3.4. Keyword Analysis

3.4.1. Keyword Co-Occurrence Analysis

The larger the graph nodes represent the keywords with more occurrences, and the frequency of keywords reflects the hot spots in the field of privacy protection

research. After combining and eliminating synonyms and irrelevant words in keywords, co-occurrence and clustering analysis, and extracting the top ten keywords in the keyword co-occurrence analysis of Chinese literature (Table 7), we can see that the hot spots of privacy protection research in China are focused on differential privacy, big data, blockchain, k-anonymity, privacy, cloud computing, location privacy, etc. (Figure 11). According to the top ten keywords in frequency in the keyword co-occurrence analysis of foreign literature (Table 8), it can be seen that foreign privacy protection research hotspots are concentrated in model, Internet, security, trust, data privacy, differential privacy, algorithm, location privacy, information privacy, etc., with high centrality of Internet (0.71), user behavior (0.71), and information privacy (0.68) (Figure 12).

Table 7. High-frequency keywords and centrality of Chinese literature.

Ranking	Keywords	Frequency	Centrality
1	Differential Privacy	249	0.03
2	Big Data	242	0.10
3	Blockchain	208	0.04
4	K-anonymous	136	0.03
5	Privacy	128	0.10
6	Cloud Computing	127	0.03
7	Location Privacy	127	0.03
8	Data Mining	124	0.01
9	Data Release	117	0.01
10	Homomorphic encryption	102	0.02

Table 8. High-frequency keywords and centrality of English literature.

Ranking	Keywords	Frequency	Centrality
1	model	599	0.02
2	internet	568	0.71
3	security	441	0.46
4	trust	435	0.48
5	Data privacy	320	0.33
6	Differential privacy	300	0.00
7	behavior	264	0.71
8	Could computing	258	0.06
9	Location privacy	250	0.07
10	Information privacy	250	0.68

through the relevant The research in the Chinese literature has focused on the regulation of data release and data sharing, and the protection of user privacy through relevant privacy policies (Figure 13). The research in English literature (Table 10) focuses on the privacy disclosure and privacy protection of online users in social media on the mobile network side, and studies privacy protection means in all aspects through information technology, such as access control, structural model, same-station encryption, searchable encryption, etc. (Figure 14). The common research hotspots in Chinese and English literature are differential privacy, location privacy, trust, information security, privacy paradox, and social networks.

Table 9. Main cluster analysis of Chinese literature keywords.

Cluster Number	Number of nodes	Year	Shear value	Clustered Keywords	Topic Overview
0	31	2015	0.986	Artificial Intelligence; Digital Economy; Data Governance; Personal Information; Digital Trade	Privacy Information Concerns in the Age of Artificial Intelligence
3	26	2013	0.909	Privacy; personal data; privacy protection; network; platform companies	Privacy protection of Internet networks
6	21	2012	0.969	Location privacy; cybersecurity; trajectory privacy; data privacy; authentication protocols	Privacy and security based on location system
7	21	2013	0.948	Privacy policy; Influencing factors; Social media; Users; Privacy and security	Social Media User Protection Policy
9	20	2012	0.909	Internet of Things; security; cloud storage; protocols; association rules	Privacy computing perspective protection measures

Table 10. Main cluster analysis results of English literature keywords.

Cluster Number	Number of nodes	Year	Shear value	Clustered Keywords	Topic Overview
0	26	2008	0.959	Online privacy; trust; privacy issues; privacy protection	Privacy concerns for online users
3	23	2016	0.928	Differential privacy; edge computing; task analytics; data extranodal awareness	Privacy computing perspective protection measures
4	22	2015	0.914	Cloud computing; access control; attribute-based encryption; searchable encryption.	Technical aspects of privacy protection
5	22	2013	0.946	Location privacy; data privacy; location-based services; homomorphic encryption	Security measures for location information
6	21	2011	0.913	Privacy paradox; Facebook; privacy computing; online social networks; self-disclosure	The Paradox of User Disclosure and Protection

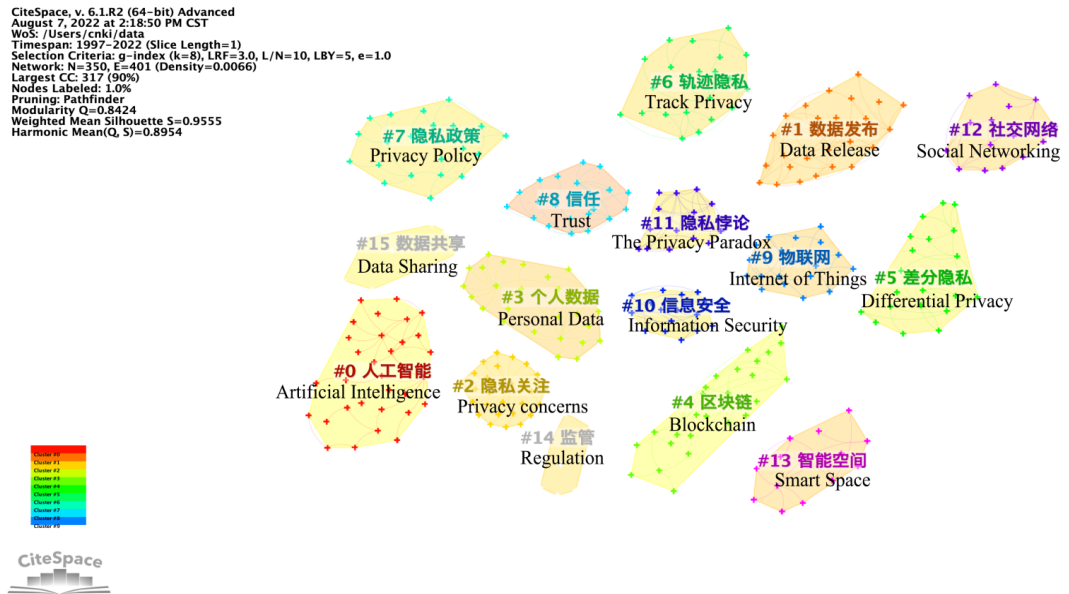


Figure 13. Clustering diagram of Chinese literature keywords based on LLR method.

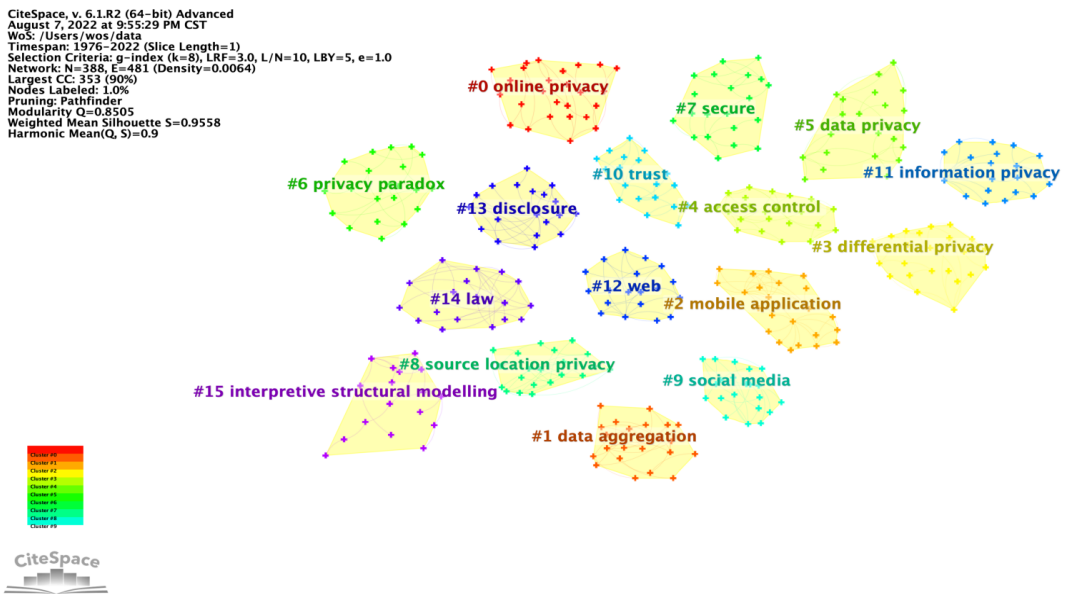


Figure 14. Clustering diagram of English literature keywords based on LLR method.

3.5. Research Frontier Analysis

Changes in research themes and hotspots in a field can be studied from keyword bursts (Wang et al., 2020), and the black horizontal line markers represent the beginning and end of keyword bursts; the longer the length of the burst, the longer the keyword hotspot lasts, the stronger the research frontier, and the higher its importance and the degree of attention it receives (Han et al., 2022). CiteSpace 6.1.R2 (64-bit) software was used to analyze the English and Chinese literature for the emergent words, and 96 English and 61 Chinese emergent words were obtained. The words privacy protection, privacy concern were hid-

den in the search formula, and the top 23 emergent words by emergent year are shown in **Figure 15** in English and **Figure 16** in Chinese.

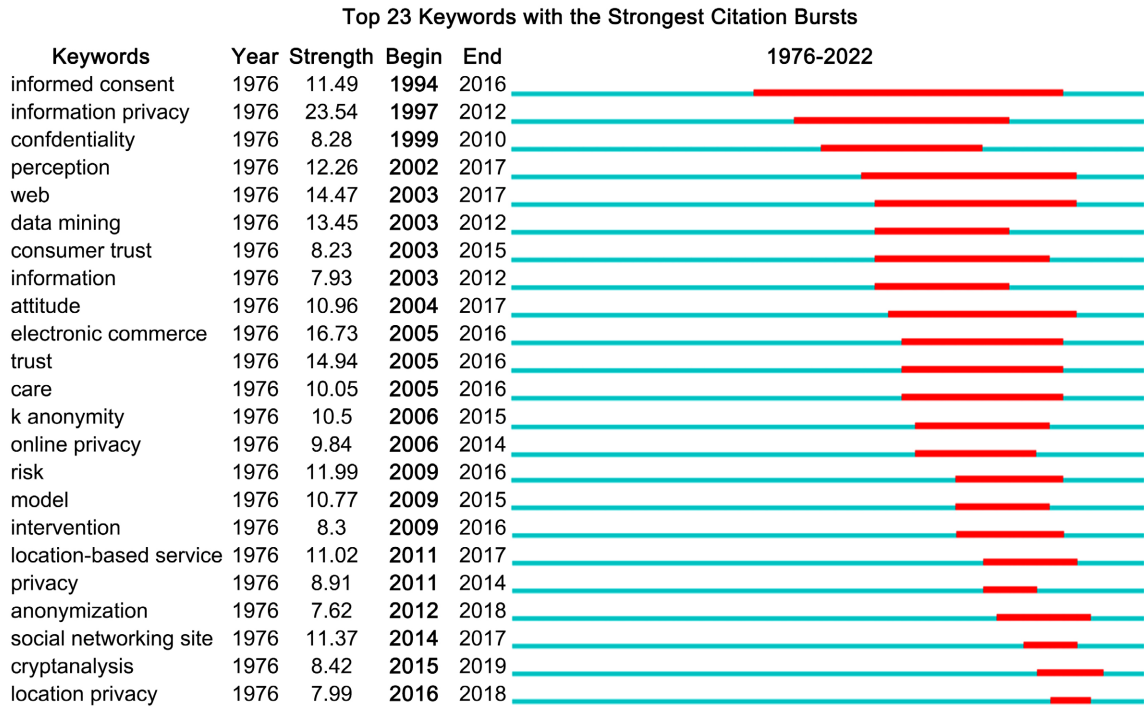


Figure 15. Top 23 emergent words in English literature.

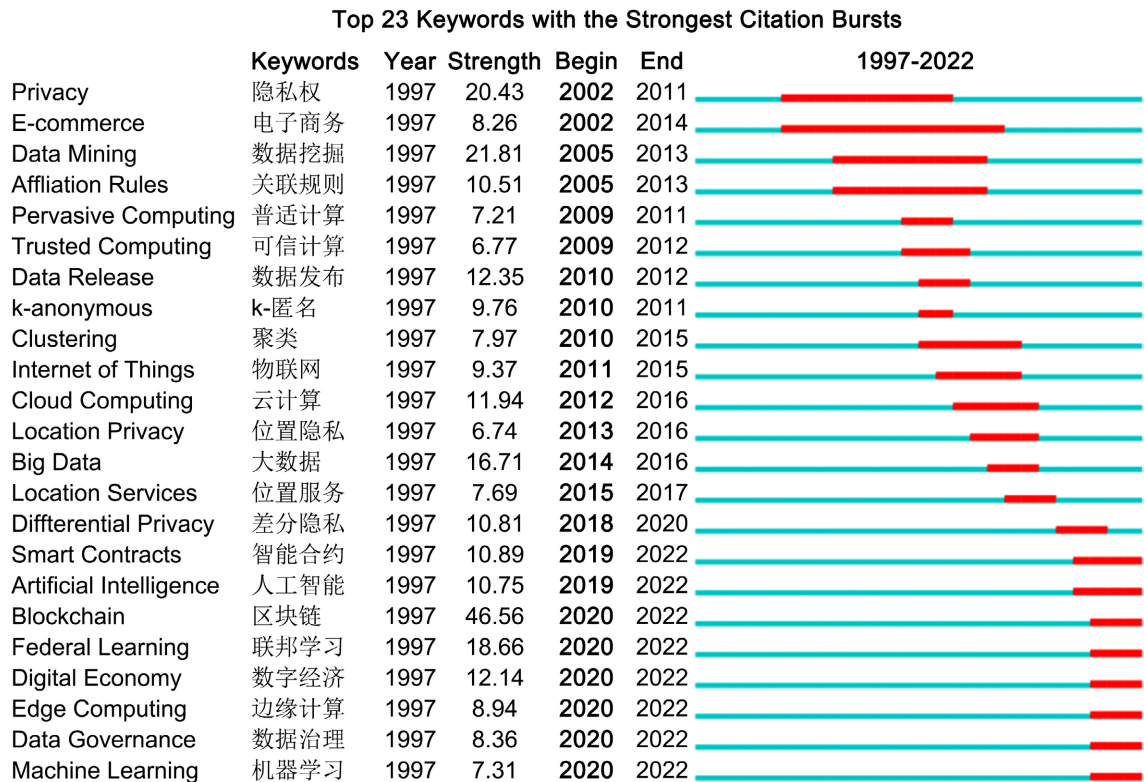


Figure 16. Top 23 emergent words in Chinese literature.

Informed consent is the keyword with the longest duration of foreign keyword emergent analysis (1994-2016), and the highest emergent intensity as of August 2022 is information privacy (23.54), followed by e-commerce (16.73) and trust (14.94). Foreign studies in recent years have focused more on encryption, location privacy and anonymization.

In China, e-commerce is the keyword with the longest duration of domestic keyword emergence analysis (2002-2014), “artificial intelligence”, “smart contract”, “Data governance” and “blockchain” have been continuing since their emergence in 2020, and are expected to become research hotspots in the field of privacy protection in the future, which researchers should pay attention to and focus on. Although domestic research on privacy protection and privacy concerns started later than foreign research, and there is a certain lag in research, the growth trend of the number of articles published in recent years is synchronized with foreign research, which indicates that domestic scholars’ enthusiasm for research in this field is increasing.

4. Discussion

4.1. Comparison of the Current Status of Domestic and International Research

After the outbreak, privacy protection and privacy concerns have exploded in international attention, and have remained above 1000 articles per year since 2019. The authors ranking 1st in English publications and the top 6 institutions are all from China, indicating the importance of Chinese research in this field, but it cannot be ignored that the centrality of the United States and the United Kingdom is very high, while that of China is only 0.05, and the level of research needs to be greatly improved. Compared with foreign countries, the number of publications in China has also been growing in recent years, with a more moderate growth trend and significantly lower than foreign countries. A visual mapping analysis of the number of publications and institutional cooperation networks by country shows that the research in the field of privacy protection and privacy concerns is in a key position in China, the United States, Australia, the United Kingdom, and South Korea.

The analysis of the co-occurrence mapping of authors and collaborating institutions shows that foreign institutions and authors have built a close collaborative network and the number of publications has increased substantially and rapidly. In contrast, the collaboration among research institutions and scholars in China is not close, and there are few large professional research teams, which makes it difficult to conduct large-scale and high-quality research in China. Moreover, a large number of Chinese scholars cooperate closely with foreign scholars and publish through foreign channels. In this regard, the author suggests that more domestic and foreign academic exchanges should be carried out, actively learn from foreign experience, and publish strong national cooperation strategies. Research institutions at the core of our country should take the lead

and take the initiative to actively develop cooperation plans with neighboring research institutions and gradually expand the scale of research groups.

4.2. Domestic and Foreign Research Content Partially Overlaps and Has Its Own Focus

Based on the analysis of keyword co-occurrence, high-frequency keywords, keyword clustering and visual graphs of emergent keywords at home and abroad, foreign countries focus on patterns, security, trust, data privacy, personalized privacy, etc., and begin to study privacy protection earlier, and have profound theoretical support, privacy protection systems and models for privacy protection. Chinese scholars' research hotspots are focused on differential privacy, big data, blockchain, location privacy, etc. Compared with foreign countries, the focus is on computer science for privacy protection. The comparison of emergent words can see the different perspectives of privacy protection at home and abroad. Foreign countries focus more on the subjective feelings of online users' privacy, such as informed consent, trust, attitude, etc., and control privacy infringement and other problems by establishing privacy protection models. Domestic is more concerned with data governance, data computing, machine learning, etc. After visualization and analysis we are able to see that domestic is starting to focus on the privacy of AI online users.

4.3. Research Features and Limitations

This research uses the current scientific and cutting-edge visualization mapping software Citespace 6.1.R2 (64-bit) to conduct the research. The mapping is a clear and intuitive way to understand the history, development and trends in this field at home and abroad, and has a high reference value. The use of domestic and foreign core data for comparative research ensures the comprehensiveness of the study. Due to the large volume of literature, only the core data of China Knowledge Network was used for Chinese literature, and the database source is relatively single. With the increase of research on privacy protection and privacy concerns, the database of Chinese literature will be enriched at a later stage and further improved for analysis in the future.

5. Conclusion

The digital age has increased the speed of the network, a multi-dimensional sensory world has arrived, and the efficiency of transmission has increased ten thousand times. Companies and governments around the world are increasingly relying on technology to collect, analyze and store personal data. If location information and trajectory data are directly released without any control, and analyzed and mined at will, it will lead to serious social problems due to the leakage of a large amount of personal privacy information. It is clear that (Sundaram & Shetty, 2022) the technological changes caused during Covid-19 affect privacy concerns and protective behaviors, privacy efficacy increases privacy

concerns and protective behaviors, and online video chat software users' privacy concerns and protective behaviors increase accordingly. Through the analysis of domestic and international data (Wan et al., 2016), location privacy protection technology gradually becomes a research hotspot and a large number of research results emerge, location data can expose a large amount of personal information, social relationships, religious beliefs, behavioral habits, health status, political tendencies, etc., so we should pay enough attention to it. With the leapfrog development of social networks, a large number of privacy leaks occur frequently at home and abroad, in addition to the technical aspects to solve the privacy protection problem, it should also be solved from the integration of various aspects such as enterprise management, consumer psychology, value perception, and privacy paradox.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Chen, Y. et al. (2015). The Methodological Function of CiteSpace Knowledge Graph. *Scientology Research*, 33, 242-253.
- China Association for Science and Technology (2018). *Blue Book on the Development of Chinese Science and Technology Journals/2017*. Science Press.
- Feng, Y. F., Yan, C., & Hu, C. P. (2020). Structural Features and Hotspot Perspective of Domestic Privacy Field Research in the Past 20 Years. *Journal of Information Resource Management*, 10, 65-74+101.
- Gu, L. P., & Fan, H. C. (2018). The Academic Field of Ten Years of Research on Online Privacy—An Analysis Based on the CiteSpace Visual Scientific Knowledge Graph (2008-2017). *Journalism and Communication Research*, 25, 57-73+127.
- Guo, W.-P., & He, D.-X. (2019). Characteristic Distribution and Hotspot Analysis of Archival Research in Chinese Universities in the Last Decade—A Bibliometric and Visual Analysis Based on CNKI Core Journals. *Archival Research*, No. 2, 25-30.
- Gutierrez, A. et al. (2019). Using Privacy Calculus Theory to Explore Entrepreneurial Directions in Mobile Location-Based Advertising: Identifying Intrusiveness as the Critical Risk Factor. *Computers in Human Behavior*, 95, 295-306.
<https://doi.org/10.1016/j.chb.2018.09.015>
- Han, L. L. et al. (2022). A Visual Analysis of Hotspots of Domestic and International Research on Fear of Childbirth Based on CiteSpace. *Nursing Research*, 36, 2093-2100.
- Huang, H. et al. (2020). An Approach to Privacy Preservation of Large-Scale Social Network Data with Weights. *Computer Research and Development*, 57, 363-377.
- Huang, H. P., Wang, K., Tang, X., & Zhang, D. J. (2019). A Differential Privacy Preservation Scheme Based on the Edge Mediation Model. *Journal of Communication*, 40, 88-97.
- Kar, A. K., & Dwivedi, Y. K. (2020). Theory Building with Big Data-Driven Research—Moving Away from the “What” towards the “Why”. *International Journal of Information Management*, 54, Article ID: 102205.

- <https://doi.org/10.1016/j.ijinfomgt.2020.102205>
- Lee, D.-J., Ahn, J.-H., & Bang, Y. (2011). Managing Consumer Privacy Concerns in Personalization: A Strategic Analysis of Privacy Protection. *MIS Quarterly*, *35*, 423-444. <https://doi.org/10.2307/23044050>
- Li, W.-P., Yang, J., Zhang, J.-P., & Yin, G.-S. (2015). Personalized Trajectory Privacy Protection Algorithm Based on CCA. *Journal of Jilin University (Engineering Edition)*, *45*, 630-638.
- Li, J., & Chen, C. M. (2017). *CiteSpace Technology Text Mining and Visualization* (2nd ed.). Capital University of Economics and Business Press. <http://www.yunzhan365.com/basic/91432764.html>
- Li, Y. P., Cao, Q., Zhang, K., & Ren, F. (2021). A Secure Index Resisting Keyword Privacy Leakage from Access and Search Patterns in Searchable Encryption. *Journal of Systems Architecture*, *115*, Article ID: 102006. <https://doi.org/10.1016/j.sysarc.2021.102006>
- Liu, T. T., & Lai, Z. S. (2022). From Non-Player Characters to Othered Participants: Chinese Women's Gaming Experience in the "Free" Digital Market. *Information, Communication & Society*, *25*, 376-394. <https://doi.org/10.1080/1369118X.2020.1791217>
- Minsky, N. (1976). Intentional Resolution of Privacy Protection in Database Systems. *Communications of the ACM*, *19*, 148-159. <https://doi.org/10.1145/360018.360027>
- Ni, W. W., Chen, G., Chong, Z. H., & Wu, Y. J. (2012). A Study on Clustering-Oriented Data Hiding and Publishing. *Computer Research and Development*, *49*, 1095-1104.
- Oberg, C., & Alexander, A. T. (2019). The Openness of Open Innovation in Ecosystems—Integrating Innovation and Management Literature on Knowledge Linkages. *Journal of Innovation & Knowledge*, *4*, 211-218. <https://doi.org/10.1016/j.jik.2017.10.005>
- Ozdemir, Z. D., Jeff Smith, H., & Benamati, J. H. (2017). Antecedents and Outcomes of Information Privacy Concerns in a Peer Context: an Exploratory Study. *European Journal of Information Systems*, *26*, 642-660. <https://doi.org/10.1057/s41303-017-0056-z>
- Rong, H. G., Ma, Y. X., Du, M. J., & Fei, Y. T. (2022). Hotspots and Frontier Analysis of Chinese Herbal Medicine Management Research Based on CiteSpace Visualization Mapping in China. *Chinese Herbal Medicine*, *53*, 4075-4083.
- Shu, S., & Liu, Y. (2021). Looking Back to Move Forward: A Bibliometric Analysis of Consumer Privacy Research. *Journal of Theoretical and Applied Electronic Commerce Research*, *16*, 727-747. <https://doi.org/10.3390/jtaer16040042>
- Sundaram, R., & Shetty, S. (2022). Privacy Concerns and Protection Behavior during the Covid-19 Pandemic. *Problems and Perspectives in Management*, *20*, 57-70. [https://doi.org/10.21511/ppm.20\(2\).2022.06](https://doi.org/10.21511/ppm.20(2).2022.06)
- Wan, S. et al. (2016). Advances in Location Privacy Protection Technology Research. *Journal of Communication*, *37*, 124-141.
- Wang, H. (2021). A Study on the Rise of Personal Privacy Protection Awareness in the Era of Big Data—A Visual Knowledge Graph Analysis Based on CiteSpace. *Henan Science and Technology*, *40*, 11-14.
- Wang, Z. Y. et al. (2020). Research Progress and Development Trend of Social Media Big Data (SMBD): Knowledge Mapping Analysis Based on CiteSpace. *ISPRS International Journal of Geo-Information*, *9*, Article No. 632. <https://doi.org/10.3390/ijgi9110632>
- Xu, J. H., Hou, W. P., & Hu, S. M. (2020). Historical Changes and Frontier Hotspots in a Century of Privacy Research in English Academia—A CiteSpace Analysis Based on the Web of Science Database. *Governance Studies*, *36*, 109-122.

- Xu, Z. Q., Xiang, D., & He, J. L. (2021). Data Privacy Protection in News Crowdfunding in the Era of Artificial Intelligence. *Journal of Global Information Management*, 30, 1-17. <https://doi.org/10.4018/JGIM.286760>
- Yang, J., Zhao, J. S., & Zhang, J. P. (2013). A Privacy Preserving Approach for High-Dimensional Data Mining. *Journal of Electronics*, 41, 2187-2192.
- Zhang, L., Ma, C. G., Yang, S. T., & Li, Z. P. (2017). An Approach to Location Privacy Preservation with Indistinguishable Association Probabilities. *Journal of Communication*, 38, 37-49.
- Zhang, Y. F., & Xu, H. X. (2022). Accountable Monero System with Privacy Protection. *Security and Communication Networks*, 2022, Article ID: 7746341. <https://doi.org/10.1155/2022/7746341>
- Zhu, S., & Yang, G. (2022). PCA Differential Privacy Data Publishing Algorithm in Transversal Federated Learning. *Computer Applications Research*, 39, 236-239+248.
- Zhu, W. (2013). A Survey and Analysis of Domestic and Foreign Social Network Privacy Protection Research. *Information Technology Research*, 39, 5-9.