[1, *]Rolly R. Tang
[1] Jae Kyu Lee
[1] Shan Liu

Taxonomy of the GDPR-based Privacy Research by Scientometric Analysis

Journal of Electrical Systems

Abstract: - Will General Data Protection Regulation (GDPR) be adopted globally in business? The GDPR was approved in the European Union (EU) in April 2016 and officially put into effect in May 2018, thus the research in this field has an obvious upward trend. The development of GDPR is aimed at adapting to new trends, conducting scientific econometric analysis in the fields of privacy and GDPR, and analyzing and visualizing emerging trends. First, summarizing the privacy and GDPR studies publicly published between 1995 and 2023 through statistical analysis of terminology categories and high-yield journals. Then, understand the overall research status of privacy rights and GDPR from the perspectives of author, journal, literature co citation analysis, and collaborative networks. Finally, based on keyword analysis and literature co citation cluster analysis, a knowledge graph was constructed that includes knowledge domains, evolutionary trends, and future research directions. As a globally influential regulation, GDPR emphasizes the protection and lawful processing of personal data, which is of great significance for protecting personal data privacy and enhancing data security.

Keywords: GDPR; Privacy; Globally; Cluster analysis.

I. INTRODUCTION

Scientometrics is a crucial method to explore the scientific research rules, identify research trends, and evaluate the development of the field [1,2]. In this paper, scientometrics analysis is performed in the privacy and GDPR (General Data Protection Regulation) domain and software named CiteSpace is utilized to analyze and visualize the emerging trends.

By using the CiteSpace and reviewing privacy and GDPR research published between 1995 and 2023, it is possible to display the evolution of a knowledge domain on a network map and to identify research frontiers. Four major questions for the body of privacy and GDPR literature are advanced: (1) What is the basic situation of term classification, journals, authors, institutions? (2) What are the citation status and influence of references, journals and authors in co-citation analysis? (3) How prominent are individual authors, institutions and countries in the corresponding collaboration network? (4) What research phases and opportunities for future research seem promising?

Accordingly, three main objectives of this study are: (1) summarize the privacy and GDPR research published during 1995-2023, by statistical analysis term categories, high-yield journals; (2) understand the overall research status for privacy and GDPR from the perspective of author, journal, reference co-citation analysis and collaboration network; (3) based upon keywords analysis and reference co-citation cluster analysis, a new integrated, holistic knowledge map that includes knowledge domains, evolutionary trends, and future research directions.

A taxonomy of information privacy which includes individual, group [3], and organizational privacy, most privacy studies in the information systems field have been conducted at the individual level of analysis, consider more than one level of analysis and beyond the individual level of analysis when necessary [4]. Privacy research may focus on organizational level [5] and individual level [6-8] perspective, while privacy and GDPR may focus more on policy and law aspects, as it is written by lawyers and policy-makers [9].

Existing research concentrated on protective approach, while few preventive measures were proposed [10], the current cybersecurity systems mainly depend on the Receiver's Responsibility Paradigm that aims the self-defensive protection of each individual system, to overcome such a limitation fundamentally, a complementary paradigm of Preventive Cybersecurity was proposed which emphasizes the importance of "Origin and Deliverer Responsibilities" [11-14]. However, the adoption of preventive cybersecurity measures are resisted because they may infringe the privacy and freedom expression of innocent netizens. The protection of privacy rights has become a global issue as European Union (EU) adopted the GDPR.

[1] School of Management, Xi'an Jiaotong University, No.28, Xianning West Road, Xi'an and 710049, China
*Corresponding author: Rolly R. Tang

II. METHODOLOGY

This section gives our methodology of scientometric analysis for privacy and GDPR, shows the collection of empirical data. The data used to analysis in our research is downloaded from WoS, and the search strategy followed is below:

(1) Themes = ("privacy" OR "information privacy" OR "the privacy" OR "personal secrets" AND "GDPR" OR "General Data Protection Regulation");

As for the definition of privacy, no single definition can be workable, but rather that there are multiple forms of privacy, scholars have suggested that privacy is one's ability to control his own information [4,18]. In matters of GDPR, it is defined as a milestone in convergence for cybersecurity and compliance, proposed new legal requirements for privacy management in the field of Information Systems.

(2) Database = "SCI-E, SSCI, CPCI-S, CPCI-SSH and ESCI";

(3) Timespan = "1995-2023";

(4) Document types = "Article" or "Editorial Material" or "Proceedings Paper" or "Review";

(5) Literature type = "English";

1798 publications are retrieved, and finally 1706 publications are valid data, which were downloaded on Dec 31st, 2023.

III. IMPACT OF GDPR BASED RESEARCH IN PRIVACY RESEARCH

A. *Dynamic Trend in Number of Publications*

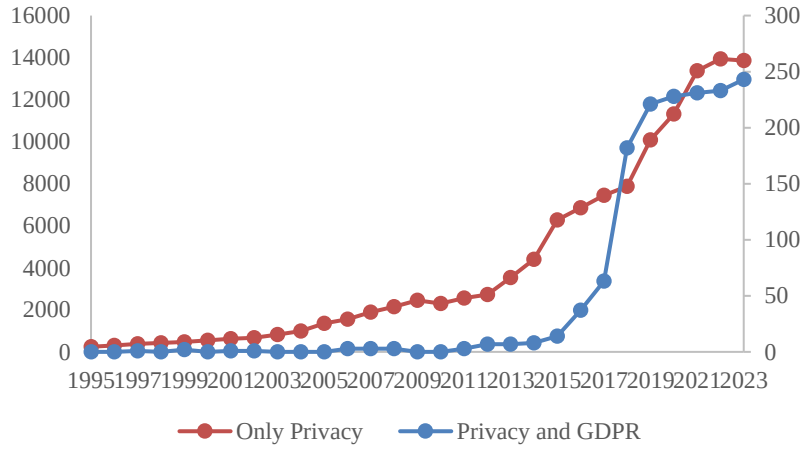

Figure 1. The dynamic trend by annual number of publications

Only Privacy shows growth trend from 1995 to 2023, in the meantime, Fig. 1 shows that the number of publications in privacy and GDPR is increasing in the past 30 years, from 1 publication in 1997 to 243 publications in 2023, with rapid growth in 2016-2023. GDPR was enacted in 2016 and put into effect in 2018, thus the research in this field has an obvious upward trend.

B. *Topics and Network of Keywords*

EU has been a pioneer in privacy and data protection issues for decades. It can be found that the high frequency keyword list mostly refers to the main terms related to privacy and GDPR such as Personal Data, Security, Data Protection, Law Enforcement, Trust, Information Privacy [4,15-17], Anonymization [6,18], Internet of Things [19,20], etc., which implies the research issues of privacy and GDPR. Keywords such as Law Enforcement, Trust, Surveillance implies the research issues (topic) of privacy, these issues may occur in various platform contexts.

In the meantime, the main technologies and tools involved in privacy and GDPR are Big data [21], Blockchain [22-24], Machine Learning, Artificial Intelligence [21], Cloud Computing [25,26], Accountability [24,27], etc.

Betweenness Centrality (BC) is an indicator to measure the importance of nodes in the network, which indicates the extent to which a node is a 'bridge' between other nodes in the network graph, nodes more than 0.1 are called key nodes, which are used to discover and measure the importance of literature.

The antecedent of GDPR is the Computer Data Protection Law enacted by EU in 1995. In the first stage (1995-2008), the main high-frequency words are Privacy and Data Protection, the BC values are 0.18 and 0.2, respectively, which are higher than 0.1, indicating that Privacy and Data Protection are research hotspot. Consent is the highest value of BC among all keywords, on account of in the definition of user data in GDPR, any collection of PII (Personal Identifier Information) must explicitly ask whether the user agrees or not.

The European Parliament proposed to reform EU Data Protection Regulations in January 2012, they agreed to formulate new EU Data Protection Regulations in December 2015. In the second stage (2009-2015), the BC (0.12) of GDPR in 2012 was the highest, and the frequency (202) of GDPR in 2015 was the highest. Ann Cavoukian initially developed the 'Privacy by Design', GDPR introduces such data protection mechanism [10], making the external entities have to accept the EU data governance concept in the process of business compliance, which has a significant impact on the business operation mode of global data processing entities.

In the third stage (2016-2023), the European Parliament adopted the GDPR in April 2016, the EU directive on citizens' right to data protection in law enforcement came into effect in May 2016, the opinions on the guidelines of leading regulatory bodies had been closed in February 2017, the final guidelines were pre-implemented for assessment, certification and punishment in April 2017, before the GDPR came into force, a lot of exploration has been done on transparency and accountability [28].

GDPR was fully carried out on May 25, 2018, which can be said to be the data privacy protection act with the highest level of data protection. After the GDPR takes effect, the privacy and protection of personal data will be more transparent and operational. The term 'Accountability' [24,27] is already a popular word that implies the 'Responsibility' in the Bright Internet, which is proposed the following five design principles: Origin Responsibility, Deliverer Responsibility, Identifiable Anonymity, Global Collaboration, and Privacy Protection, ensuring that the privacy protection and freedom of expression of innocent netizens [11,12,14], this is extremely consistent with Anonymization and Accountability above.

C. Contexts of Privacy Research

1) Co-occurrence Network Analysis

The keywords co-occurrence map is clustered into irregular regions, each region has its corresponding label, and each cluster is composed of several closely related words, which are all keywords in the co-occurrence network.

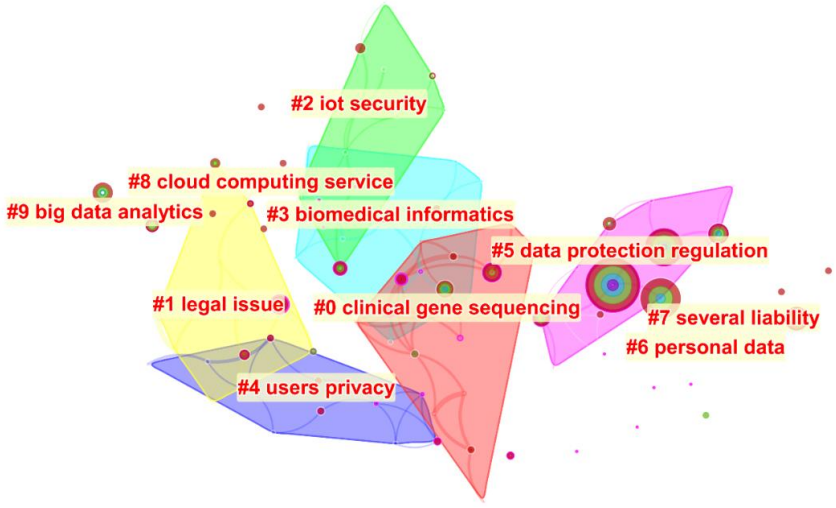

Figure 2. Keyword clusters network

The research clusters and the relative importance rank based on keyword is shown in Fig. 2, consequently, cluster IDs with the newest group (Purplish red) are cluster #5 "Data Protection Regulation", #6 "Personal Data" and #7 "Several Liability" [21] with the largest sized cluster in 2018, 2019. The size of the node in the figure represents the frequency of keyword, the larger the node, the greater the frequency of the keyword and the greater the relevance to the topic, it is obvious that recent development in privacy and GDPR research has centered on clusters #5, #6 and #7, as shown by the Fig. 2.

From the keyword cluster network, cluster #1 "Legal Issue" (yellow) and cluster #4 "Users Privacy" (blue purple) are also hot topics in the mainstream, this is consistent with the previous documents co-citation research. Using the keyword co-occurrence network, the subject structure in the dataset can be clearly displayed, Figs. 2.

also suggest that cluster #2 “IoT Security” [19,20], cluster #8 “Cloud Computing Service” [25,26] and cluster #9 “Big Data Analytics” [21] are among the most popular techniques of that cluster. By and large, the keywords in Figs. 2. seem to classify the privacy issues in the context of IoT security, in the context of cloud computing, and in the context of big data analytics.

2) Documents Co-citation Cluster Analysis

There are 585 nodes and 1495 edges, in the network of documents co-citation, each node represented one document, the links of the connected nodes represented co-citation relationships, larger nodes indicate articles cited by numerous different scholars. Within the network of document co-citation, 53 articles (7.74 percent) were published in Computer Law & Security Review (England, impact factor 2.9), it is the source journal with most publications on privacy and GDPR topics. The majority of them are conference proceedings, they are fundamental to privacy and GDPR research publishing, but it is difficult for readers to visualize their influence and citation status by only using such few journal articles, so there is a potential research topic in such field.

The nodes with both high BC and high frequency characteristics are the key literature in this field, and also the key literature in this period, representing the hot topic and frontier of this period. Figure 3 shows the citation knowledge network and clustering results of privacy and GDPR. It can be seen that the network has ten clusters of a combined model, revoking consent [29], big data [21], GDPR readiness, privacy design strategies [9,10], two-fold shift, artificial intelligence [21], intervenability requirement, processing children, data protection compliance regulation [30], and information technology sector in Fig. 3.

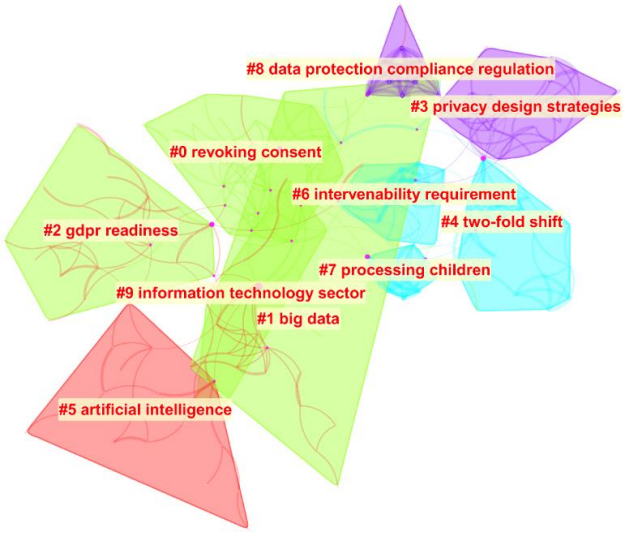

Figure 3. Document co-citation clusters network

From the clustering results, we can find that the earliest cluster is artificial intelligence, which illustrates the importance of artificial intelligence in this field. Big data and GDPR readiness are the contents of early attention, data protection compliance regulation [30] and privacy design strategies [9,10] are more concerned in recent years.

The cluster details show the cluster names obtained by three methods (LSI, Log-Likelihood Ratio and Mutual Information), which reflect the research frontier fields, and the citation literature shows the research frontier. The research results of Dr. Chaomei Chen show that the clustering identification of LLR word selection algorithm is relatively representative and comprehensive, because it is closest to manual tags [31], in addition, LLR can generate high-quality clusters with intra-class similarity and low inter-class similarity [32].

Table 1. Largest clusters of co-cited documents

Cluster ID	Size	Silhouette	Label (LLR)
8	10	0.966	data protection compliance regulation
3	22	0.96	privacy design strategies
5	17	0.956	challenging algorithmic profiling
9	9	0.954	information technology sector
7	10	0.947	processing children
1	27	0.946	big data

4	20	0.939	two-fold shift
6	16	0.929	intervenability requirement
2	26	0.908	GDPR readiness
0	28	0.847	revoking consent

In Table 1, Silhouette is a measure of cluster's homogeneity, and the closer its value is to 1, the more homogeneous it is, all silhouettes of ten clusters are greater than 0.8, which means the clustering results are reliable. The cluster name is extracted by Log-Likelihood Ratio (LLR) method, the frontier fields are data protection compliance regulation [30] and privacy design strategies [9,10], Size represents the number of articles in a cluster, and there are 10 articles in the cluster (#8) and 22 articles in the cluster (#3). Mean (Year) represents the average year of publications in a cluster, and it is used to evaluate the average time when the cluster appears, the cyberlaws GDPR was released its initial proposal in Jan, 2012 in the cluster (#8, #3), adopted in Apr, 2016 in the cluster (#5, #9).

IV. Is GDPR Global Trend?

A. GDPR was Born in EU, and Researched in EU

The 6 founding members of EU are Germany (255), Netherlands (154), Italy (165), Belgium (102), France (87), Luxembourg (44), England (214) may be regarded as EU which GDPR was developed, it ranks third in the world which officially left EU on January 31, 2020.

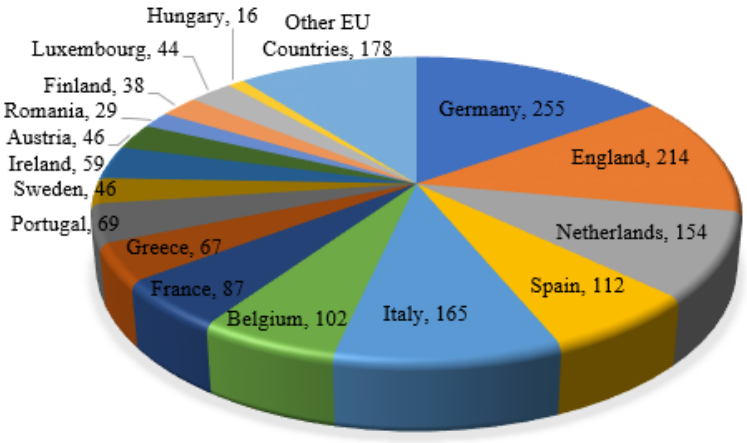

Figure 4. Distribution of publications in EU

Fig. 4 shows that other EU Member States are Spain (112), Portugal (69), Greece (67), Ireland (59), Sweden (46), Austria (46), Finland (38), Romania (29), Hungary (16), respectively.

B. Wide Publications from USA and Other Non-EU Countries

In contrast with EU distribution, the dynamic trend of propagated to outside of Europe is as follows, the USA ranks first in the world, Table 4 shows that other Non-EU Member States are Australia (60), Norway (55), Switzerland (53), Canada (49), Scotland (31), Brazil (25), South Korea (24), respectively.

Table 2. Distribution of publications in Non-EU

Num	Non-EU Country	Numbers of Publications	% of total
1	USA	237	13.18
2	People's Republic of China	78	4.34
3	Australia	60	3.34
4	Norway	55	3.06
5	Switzerland	53	2.95
6	Canada	49	2.73
7	Scotland	31	1.72
8	Taiwan	26	1.45
9	Brazil	25	1.39

10	South Korea	24	1.34

Meanwhile, People's Republic of China (Mainland China, 78) and Taiwan (26) contributed to the study of privacy and GDPR.

C. Country Cooperation Network

The cooperation network of country is also the national co-occurrence map, the size of nodes in the network reflects the amount of papers published by the country. It is drawn according to the circumstance of cooperation among countries in the cited literature, the appearance of two countries in the same article is regarded as a cooperation, which is mainly based on the co-occurrence frequency matrix of countries, discussing which countries have cooperation in similar topics.

The countries cooperative network for privacy and GDPR has 63 nodes and 207 edges, as shown in Fig. 5. In-network, 23 countries were identified by relative contribution (more than 10 articles) to privacy and GDPR area, we can find that Germany, England and USA are the largest contributors to the countries cooperation network in such field.

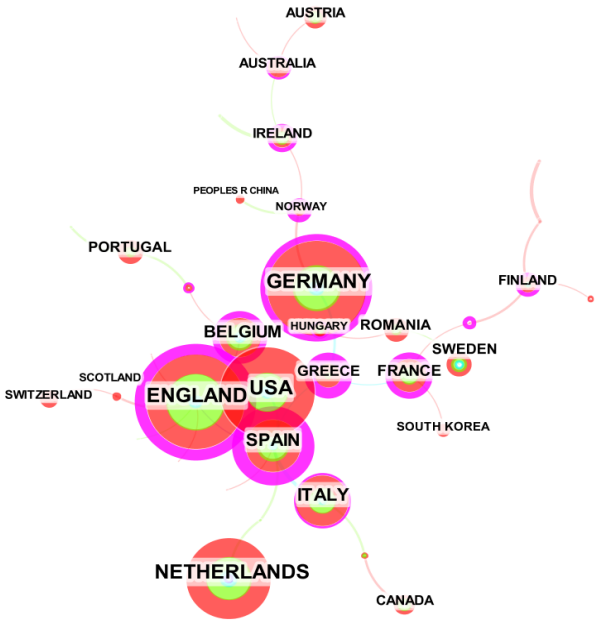

Figure 5. The network of countries for privacy and GDPR research

High frequency nodes represent highly cited literature, the node of high Betweenness Centrality represents the literature that forms the co-citation relationship with multiple literatures, from Fig. 5, the top 10 countries are the USA (88 articles, 0.18), Germany (85 articles, 0.15), England (82 articles, 0.29), Netherlands (68 articles, 0.09), Spain (53 articles, 0.12), Italy (47 articles, 0.1), Belgium (38 articles, 0.04), France (29 articles, 0.08), Greece (25, 0.05), Australia (25, 0.09). These countries are the cooperation between countries in publishing articles, they are core nodes establishing links with other nodes in the countries' collaboration network.

V. PRIVACY RESEARCH IS MULTIDISCIPLINARY NATURE

A. Disciplines Nature (CS, Law, etc.)

Privacy is interdisciplinary research, [4,16,33-35]. From the perspective of term categories distribution, as shown in Fig. 6, privacy and GDPR is a multidisciplinary research field, which mainly involving Computer Science Information Systems (23.20%), Computer Science Theory Methods (21.86%) and Law (21.56% of the total).

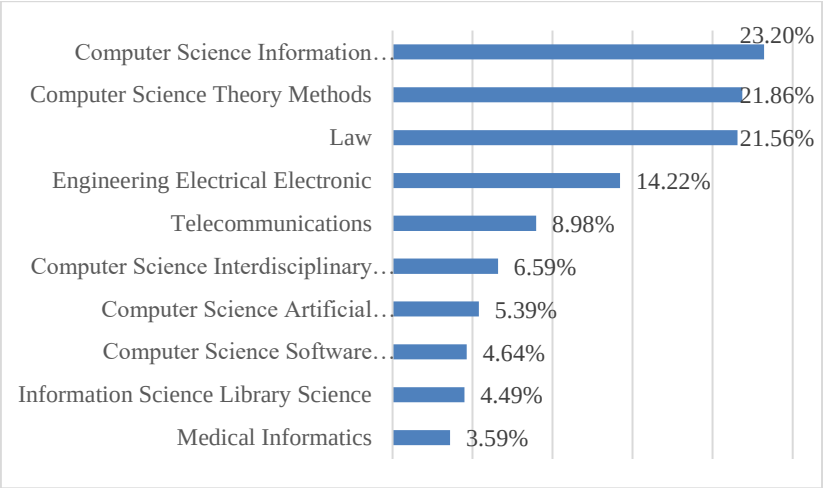

Figure 6. Distribution of the Term categories with the top 10 literature

Set spams originating from Korea as the treatment group and spams originating from other countries as the control group, discussed the Anti-spam legislation with difference-in-difference (DID) regression Analysis, examined whether the enacted antispam policy based on the 'opt-in' approach in South Korea can effectively decrease the number of spam messages originating from South Korea [36]. Examined the impact of the Real Name Verification Law policy which South Korean government implemented on privacy and anonymous issues, explored the effects of the law with real world dataset in terms of privacy and anonymity [37].

The node size reflects the frequency of the research field, the category contributed to privacy and GDPR research consisted of 162 nodes and 474 links (presented in Fig. 7). The top WoS categories are Computer Science (Theory & Methods, Information Systems, Interdisciplinary Applications, Artificial Intelligence, Software Engineering), Government & Law, Law, Engineering (Electrical & Electronic), Telecommunications, Business & Economics, the co-occurrence analysis of the privacy and GDPR field is consistent with the current research hotspots and research frontiers.

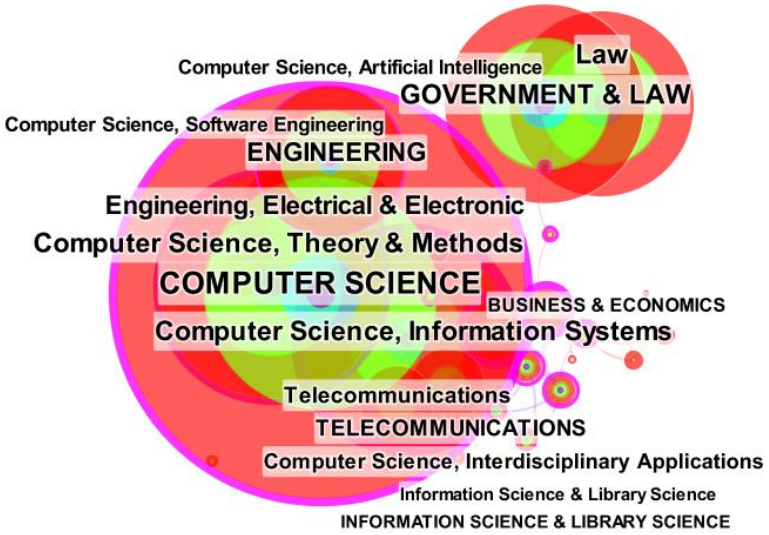

Figure 7. Visualization of the category of privacy and GDPR

Reviewed the literature regarding privacy in the field of information systems, highlighted the following: "Information systems research should focus more on design and action with an emphasis on building actual implantable tools to protect information privacy." [4]. Future studies on privacy in the information systems domain should address the design science perspective, which aims to build the tools and technologies regarding various aspects of information privacy [15].

B. Co-citation between Different Disciplines: between Journals and Authors

By using the journals cited to generate a network of co-cited journals, demonstrating 626 nodes and 1694 links, the co-citation network at journal level is shown in Fig. 8, Lecture Notes in Computer Science is most

prominent with 153 co-citations, followed by Computer Law & Security Review (141), International Data Privacy Law (112), Communications of the ACM (91), and Harvard Law Review (60).

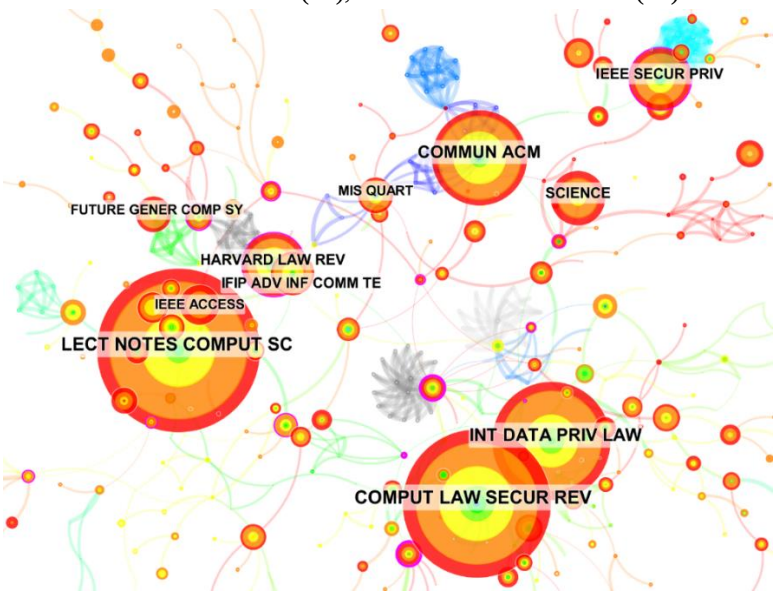

Figure 8. Network of journals' co-citation

Other high co-citation journals were identified from the network by setting the frequency threshold of 34, they are IEEE Security & Privacy (59, USA), IEEE Access (39, USA), Future Generation Computer Systems (35, Netherlands), conference proceedings IFIP Advances in Information and Communication Technology (42), there are also some top journal articles worth focusing on continuous follow-up, such as Science (52) and Management Information Systems Quarterly (34).

Fig. 9 shows the author co-citation network that contributes to privacy and GDPR, which contains 627 nodes and 1,771 co-citation links. In the author co-citation network, the node frequency of Ann Cavoukian is the largest (frequency=64, 2014), as a former Ontario privacy commissioner, Ann Cavoukian is one of Canada's foremost privacy experts and a proponent for ensuring the privacy rights of citizens. From the first time of document in 2014 that GDPR adopted, all the authors have carried out the related study early and made a consistent contribution.

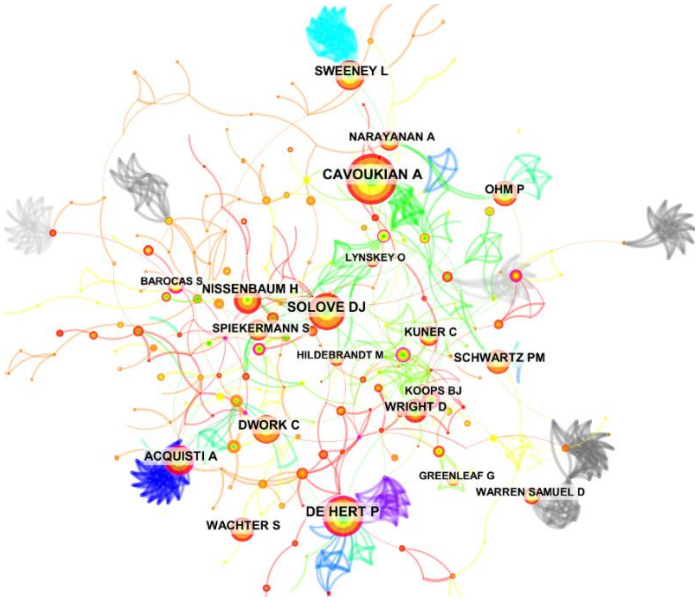

Figure 9. Network of author co-citation

In the following content, we will analyze the five privacy scholars Paul De Hert, Daniel J. Solove, Paul M. Schwartz, Alessandro Acquisti, and Latanya Sweeney, respectively. They have different contribution types, Paul De Hert, Alessandro Acquisti, and Latanya Sweeney have contributed to academic journal articles, while Daniel J. Solove and Paul M. Schwartz have contributed to books.

First, Paul De Hert (frequency=52, BC=0.16) comes from Free University of Brussels (Belgium) and Tilburg University (Netherlands), his most of the research achievements concerning privacy and GDPR were published in Computer Law & Security Review. He made continuous tracking research on privacy cyberlaws regulation GDPR [38,39], which released its initial proposal in Jan, 2012, in the following research, he elaborated the new right to data portability by empowering approach [40]. His research team also discussed the privacy authentication under the GDPR in data protection [39], explored GDPR and the NIS Directive about the processing of personal data in network and information systems [41], we should pay more attention to data protection principles enacted in Article 5 of GDPR to ensure legal certainty [42].

Second, Daniel J. Solove (frequency=48, BC=0.11), an internationally-known expert in privacy law, he founded TeachPrivacy, a company providing privacy and data security training. As an authority on information privacy law, developed a taxonomy to understand privacy violations [43] and proposed several ways privacy law can grapple with the consent dilemma and move beyond relying too heavily on privacy self-management [44]. He has written numerous books including Nothing to Hide: The False Tradeoff Between Privacy and Security [45], Understanding Privacy (D. J. Solove, 2008), The Future of Reputation: Gossip, Rumor, and Privacy on the Internet [46], and The Digital Person: Technology and Privacy in the Information Age [47].

Third, as a leading international expert on information privacy law, Paul M. Schwartz (frequency=39, BC=0.06) and Daniel J Solove co-authored several textbooks including Information Privacy Law, Privacy Law Fundamentals [48], Privacy and the Media, Privacy, Law Enforcement, and National Security, Consumer Privacy and Data Protection, and Privacy, Information, and Technology.

Furthermore, Alessandro Acquisti (frequency=41, BC=0.11) have been published a great many of top journal articles on privacy research across multiple disciplines, such as Science [33], Management Information Systems Quarterly [6], Management Science [34,49], Information Systems Research [17], Journal of Consumer Research [50], Marketing Science, Journal of Marketing Research [51], Journal of the Association for Information Systems [52], Journal of Economic Literature [53].

Finally, Latanya Sweeney (frequency=38, BC=0.10), the founding Director of the Data Privacy Lab in IQSS (Institute for Quantitative Social Science) at Harvard, her one focus area is data privacy, proposed 11-Anonymity Protocol, Plus Protocol, Enclave Protocol, Standardized Protocol such four “best practice” protocols to protect the data information privacy of individuals, in addition associate real names to “anonymized” data records [54], with regard to ‘k-Anonymity’, she also explored in the early stage [55,56]. Her most of the accomplishments concerning privacy were published in Technology Science, she currently holds the privacy and security seat of the Federal Health Information Technology Policy Committee, thus produced her contributions to Health Data Privacy, such as [57]. With the latest perspective, ten privacy rights in the preventive cybersecurity measures based on Bright Internet which the GDPR require protect [13,58].

C. Networked Coauthors and Institutions of Multidisciplinary

1) Networked Institutions of Multidisciplinary

The institutional cooperative network in privacy and GDPR has 240 nodes and 128 edges. Tilburg University is the leading contributors to the cooperation in this field and have published the most articles. 13 institutions which have more than 5 connections are listed: Radboud University Nijmegen (9, Netherlands), University of Amsterdam (8, Netherlands), Katholieke University Leuven (8, Belgium), University of Luxembourg (8, Luxembourg), The University of Edinburgh (7, England), Karlstad University (6, Sweden).

The six countries with frequency=5 are University of Oxford (England), Vrije Universiteit Brussel (Belgium), Universidad Politecnica de Madrid (Spain), Leiden University (Netherlands), Delft University of Technology (Netherlands), University of Nottingham (England), we find that the cross-border cooperation network is widespread.

2) Networked Coauthors of Multidisciplinary

The author's cooperative network shows the cooperation of all authors in the field of privacy and GDPR. There are 296 nodes and 280 edges in the authors’ cooperation network, we extracted the three largest cooperative networks from the authors’ cooperation network. The node size is proportional to the number of author’s publications, and the connection between nodes represents the author's cooperative relationship. The thickness of the connection represents the strength of the cooperation between the authors.

The largest cooperative network is that Wouter Joosen, Pierre Dewitte, Peggy Valcke, Davy Preuveneers, Kim Wuyts, Dimitrl Van Landuyt and other authors constitute a research cluster. Anotherr network is that Cesare Bartolini, Monica Palmirani, Arianna Rossi and Michele Martoni and other authors constitute the research

cluster, while Guillaume Scerri, Nicolas Anciaux, Lulian Sandu Popa and Luc Bouganim and other authors constitute a research cluster. Although there are many participants, there are more networks of less than 4 partners in the cooperation network, indicating that the cooperation in the field of privacy and GDPR is inadequate.

On the whole, most of the research achievement in the field of privacy and GDPR are still concentrated in a small range of scholars, and most researchers have only published a few papers. As a new field, the core author group of privacy and GDPR research is still in the formation stage, with more scholars' attention on privacy and GDPR research, the number and scope of the core author group will be further expanded in the future, which will play a guiding role in the research of privacy and GDPR, and constantly push the research to a new level.

VI. Discussions and Future Directions

GDPR has started as a privacy policy regulation of EU. The impact of GDPR is not only the whole EU companies, but also all international companies who operate their business in EU too. For the successful adoption of GDPR, four critical dimensions of propagation are studies with respect to the impact GDPR to privacy research, geographical expansion, expansion of focus from policy to business, and expansion to multi-disciplines. To measure the trends of propagation at this embryonic state, we study the statistics of academic publications in these dimensions using scientometrics analysis. We found that future directions are still needed to enrich this field as follows.

(1) Emerging needs of Preventive Cybersecurity Paradigm with global perspective. Nowadays humanity has entered the digital age of interconnectivity, can we still use the legal principles of the pre-network era to solve the ubiquitous network (networked, digitized, intelligent) issues of interconnectivity? In this situation, the understanding of the rules related to the protection of subject rights in personal information processing should shift from the subject's own defense norms to more suitable behavioral norms. The current network security system can only adopt the "receiver accountability paradigm" of self-defense, which cannot eliminate cyber-crime from the source. In order to effectively alleviate the current increasingly severe situation of cross-border cybercrime, we intends to eliminate the threat of cybercrime global at the source and realize preventive cybersecurity based on Bright Internet which the GDPR require while protecting the privacy of innocent netizens according to global standards.

(2) More research is needed to evaluate the impact of GDPR on privacy regulation and data governance. The interdisciplinary nature of ten privacy rights-related issues is emphasized, it has a significant impact on the business operation mode of global data processing entities, especially the introduction of ‘Privacy by Design’ data protection mechanism in GDPR, which makes the external entities forced to accept the EU data governance concept in the process of business compliance [60]. GDPR is in the transitional phase, discussed the privacy policies from inside and outside the EU, pre-GDPR and the post-GDPR about semantic text-features analysis [61,62]. It is paradoxical that GDPR terms ‘it is not erasable’ and ‘Right to be Forgotten’ have the opposite effect with blockchain technology [23]. After the introduction of GDPR, even websites that are not bound by GDPR can increase their market concentration in network technology services, and websites are more likely to retain top suppliers [59,63], as has been seen in UTD articles published in recent years.

(3) Privacy design. The implementation of GDPR has had a profound impact on global privacy protection legislation, and multiple countries have referenced the legislative techniques of the EU's GDPR. The personal information protection concepts and value orientations contained behind GDPR may have a more profound impact on the development and regulatory direction of the global digital age. Apple CEO Tim Cook eloquently summarized the essence of privacy design: "You can't fix privacy with bolts, you need to consider privacy issues in the product development process, you must design it in.". The method of data protection design described in GDPR has fundamental flaws. It is undoubtedly that through legislation is the first and most important step to protect the privacy of users, and enterprises still need to form professional security teams, actively meet regulatory needs, and even cooperate with some laboratories and standard certification bodies in the industry to obtain security and privacy certification and self-certification compliance.

Acknowledgment

Funding Statement: This research was supported by Shaanxi Province “14th Five-Year Plan” Education Science Planning Project (SGH22Y1872) and Xi'an Science and Technology Bureau Science and Technology Plan Funds (22GXFW0109).

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] Kim, M. C.,Chen, C. M. (2015). A Scientometric Review of Emerging Trends and New Developments in Recommendation Systems. Scientometrics, 104(1), 239-263.

[2] Olawumi, T. O., Chan, D. W. M. (2018). A Scientometric Review of Global Research on Sustainability and Sustainable Development. Journal of Cleaner Production, 183, 231-250.

[3] Skinner, G., Han, S., Chang, E. (2006). An Information Privacy Taxonomy for Collaborative Environments. Information Management & Computer Security 14(4), 382-394.

[4] Bélanger, F., Crossler, R. E. (2011). Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. MIS Quarterly, 35(4), 1017-1042.

[5] Xu, F., Luo, X., Zhang, H. Y., Liu, S., Huang, W. (2019). Do Strategy and Timing in It Security Investments Matter? An Empirical Investigation of the Alignment Effect. Information Systems Frontiers, 21(5), 1069-1083.

[6] Adjerid, I., Acquisti, A., Loewenstein, G. (2018a). Beyond the Privacy Paradox: Objective Versus Relative Risk in Privacy Decision Making. MIS Quarterly, 42(2), 465-488.

[7] Malhotra, N. K., S., K. S., J., A. (2004). Internet Users' Information Privacy Concerns (Iuipc): The Construct, the Scale, and a Causal Model. Information Systems Research, 15(4), 336-355.

[8] Zhang, X., Liu, S., Chen, X., Wang, L., Gao, B. J.,Zhu, Q. (2018). Health Information Privacy Concerns, Antecedents, and Information Disclosure Intention in Online Health Communities. Information & Management, 55(4), 482-493.

[9] Tamburri, D. (2020). Design Principles for the General Data Protection Regulation (GDPR): A Formal Concept Analysis and Its Evaluation. Information Systems, 91, 101469.

[10] Cavoukian, A. (2010). Privacy by Design: The Definitive Workshop. Identity in the Information Society, 3(2), 247-251.

[11] Lee, J. K. (2015). Research Framework for Ais Grand Vision of the Bright Ict Initiative. MIS Quarterly, 39(2).

[12] Lee, J. K. (2016). Invited Commentary—Reflections on Ict-Enabled Bright Society Research. Information Systems Research, 27(1), 1-5.

[13] Lee, J. K., Chang, Y., Kwon, H. Y., Kim, B. (2020). Reconciliation of Privacy with Preventive Cybersecurity: The Bright Internet Approach. Information Systems Frontiers, 22(1), 45-57.

[14] Lee, J. K., Cho, D., Lim, G. G. (2018). Design and Validation of the Bright Internet. Journal of the Association for Information Systems, 19(2), 63-85.

[15] Pavloupaul, A. (2011). State of the Information Privacy Literature: Where Are We Now and Where Should We Go? MIS Quarterly, 35(4), 977-988.

[16] Smith, H. J., Dinev, T., Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. MIS Quarterly, 35(4), 989-1016.

[17] Tsai, J. Y., Egelman, S., Cranor, L. F., Acquisti, A. (2011). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. Information Systems Research, 22(2), 254-268.

[18] Hintze, M., El Emam, K. (2018). Comparing the Benefits of Pseudonymisation and Anonymisation under the GDPR. Journal of Data Protection & Privacy, 2(2), 145-158.

[19] Wachter, S. (2018a). The GDPR and the Internet of Things: A Three-Step Transparency Model. Law, Innovation and Technology, 10(2), 266-294.

[20] Wachter, S. (2018b). Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR. Computer Law & Security Review, 34(3), 436-449.

[21] Wachter, S., Mittelstadt, B. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai. Columbia Business Law Review, 2019(2), 494-620.

[22] Bernabe, J. B., Canovas, J. L., Hernandezramos, J. L., Moreno, R. T., Skarmeta, A. F. (2019). Privacy-Preserving Solutions for Blockchain: Review and Challenges. IEEE Access, 7, 164908-164940.

[23] Humbeeck, A. V. (2019). The Blockchain-GDPR Paradox. Journal of Data Protection & Privacy, 2(3), 208-212.

[24] Truong, N. B., Sun, K., Lee, G. M., Guo, Y. (2019). GDPR-Compliant Personal Data Management: A Blockchain-Based Solution. IEEE Transactions on Information Forensics Security, 15, 1746-1761.

[25] Krystlik, J. (2017). With GDPR, Preparation Is Everything. Computer Fraud & Security, 2017(6), 5-8.

[26] Solove, D. J., Hartzog, W. (2014). The Ftc and Privacy and Security Duties for the Cloud. 13 BNA Privacy & Security Law Report 577, Available at SSRN: https://ssrn.com/abstract=2424998.

[27] Wachter, S., Mittelstadt, B., Russell, C. (2017). Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. Harvard Journal of Law & Technology, 31(2), 841-888.

[28] Wachter, S., Mittelstadt, B., Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), 76-99.

[29] Politou, E., Alepis, E., Patsakis, C. (2018). Forgetting Personal Data and Revoking Consent under the GDPR: Challenges and Proposed Solutions. Journal of Cybersecurity, 4(1), 1-20.

[30] Zerlang, J. (2017). GDPR: A Milestone in Convergence for Cyber-Security and Compliance. Network Security, 2017(6), 8-11.
[31] Chen, C. M., Fidelia, L. S., Hou, J. H. (2010). The Structure and Dynamics of Co Citation Clusters: A Multiple Perspective Co-Citation Analysis. Journal of the Association for Information Science Technology, 61(7), 1386-1409.
[32] Fang, Y., Yin, J., Wu, B. H. (2018). Climate Change and Tourism: A Scientometric Analysis Using Citespace. Journal of Sustainable Tourism, 26(1), 108-126.
[33] Acquisti, A., Brandimarte, L., Loewenstein, G. (2015). Privacy and Human Behavior in the Age of Information. Science, 347(6221), 509-514.
[34] Adjerid, I., Acquisti, A., Loewenstein, G. (2018b). Choice Architecture, Framing, and Cascaded Privacy Choices. Management Science, 65, 2267-2290.
[35] Magi, T. J. (2011). Fourteen Reasons Privacy Matters: A Multidisciplinary Review of Scholarly Literature1. The Library Quarterly, 81(2), 187-209.
[36] Ju, J., Cho, D., Lee, J. K., Ahn, J. (2016). An Empirical Study on Anti-Spam Legislation. Thirty Seventh International Conference on Information Systems, Dublin.
[37] Cho, D. (2013). Real Name Verification Law on the Internet: A Poison or Cure for Privacy? In: Schneier B. (eds) Economics of Information Security and Privacy III. Springer, New York.
[38] De Hert, P., Papakonstantinou, V. (2012). The Proposed Data Protection Regulation Replacing Directive 95/46/Ec : A Sound System for the Protection of Individuals. Computer Law & Security Review, 28(2), 130-142.
[39] De Hert, P., Papakonstantinou, V. (2016). The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals? Computer Law & Security Review, 32(2), 179-194.
[40] De Hert, P., Papakonstantinou, V., Malgieri, G., Beslay, L.,Sanchez, I. (2018). The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services. Computer Law & Security Review, 34(2), 193-203.
[41] Markopoulou, D., Papakonstantinou, V., De Hert, P. (2019). The New Eu Cybersecurity Framework: The Nis Directive, Enisa's Role and the General Data Protection Regulation. Computer Law & Security Review, 35(6).
[42] Gonzalez, E. G., De Hert, P. (2019). Understanding the Legal Provisions That Allow Processing and Profiling of Personal Data—an Analysis of GDPR Provisions and Principles. ERA Forum, 2019(4), 597-621.
[43] Solove, D. J. (2006). A Taxonomy of Privacy. University of Pennsylvania Law Review, 154(3), 477-564.
[44] Solove, D. J. (2013). Privacy Self-Management and the Consent Dilemma. Harvard Law Review, 126(7), 1880-1903.
[45] Solove, D. J. (2011). Nothing to Hide: The False Tradeoff between Privacy and Security. Yale University Press.
[46] Solove, D. J. (2007). The Future of Reputation Gossip, Rumor, and Privacy on the Internet. Yale University Press, Available at SSRN: https://ssrn.com/abstract=2899125.
[47] Solove, D. J. (2004). The Digital Person: Technology and Privacy in the Information Age. GWU Law School Public Law Research Paper No. 121, NYU Press, Available at SSRN: https://ssrn.com/abstract=609721.
[48] Solove, D. J., Schwartz, P. M. (2019). Privacy Law Fundamentals, Fifth Edition. IAPP, ISBN 978-971-948771-948725-948772.
[49] Adjerid, I., Acquisti, A., Telang, R., Padman, R., Adler-Milstein, J. (2016). The Impact of Privacy Regulation and Technology Incentives: The Case of Health Information Exchanges. Management Science, 62(4), 1042-1063.
[50] John, L. K., Acquisti, A. ,Loewenstein, G. (2011). Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information. Journal of Consumer Research, 37(5), 858-873.
[51] Acquisti, A., John, L. K., Loewenstein, G. (2012). The Impact of Relative Standards on the Propensity to Disclose. Journal of Marketing Research, 49(2), 160-174.
[52] Oh, W., Acquisti, A., Sia, C. L. (2018). Ict Challanges and Opportunities in Building a "Bright Society". Journal of the Association for Information Systems, 19(2), 58-62.
[53] Acquisti, A., Taylor, C. R., Wagman, L. (2016). The Economics of Privacy. Journal of Economic Literature, 54(2), 442-492.
[54] Sweeney, L., Von Loewenfeldt, M., Perry, M. (2018). Saying It's Anonymous Doesn't Make It So: Re-Identifications of "Anonymized" Law School Data. Technology Science, November 13, 2018. https://techscience.org/a/2018111301.
[55] Sweeney, L. (2002a). Achieving K-Anonymity Privacy Protection Using Generalization and Suppression. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5), 571-588.
[56] Sweeney, L. (2002b). K-Anonymity: A Model for Protecting Privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5), 557-570.
[57] Yoo, Y. S., Thaler, A., Sweeney, L., Zang, J. Y. (2018). Risks to Patient Privacy: A Re-Identification of Patients in Maine and Vermont Statewide Hospital Data. Technology Science, October 09, 2018. https://techscience.org/a/2018100901.
[58] Ke, T. T., & Sudhir, K. (2023). Privacy Rights and Data Security: GDPR and Personal Data Markets. Management Science, 69(8), 4389-4412.
[59] Johnson, G. A., Shriver, S. K., & Goldberg, S. G. (2023). Privacy and Market Concentration: Intended and Unintended Consequences of the GDPR. Management Science, 69(10), 5695-5721.
[60] Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory Spillovers and Data Governance: Evidence from the GDPR. Marketing Science, 41(4), 746-768.
[61] Linden, T., Khandelwal, R., Harkous, H., Fawaz, K. (2020). The Privacy Policy Landscape after the GDPR. Proceedings on Privacy Enhancing Technologies, 2020(1), 47-64.

[62] Jia, J., **, G. Z., & Wagman, L. (2021). The Short-Run Effects of the General Data Protection Regulation on Technology Venture Investment. Marketing Science, 40(4), 661-684.
[63] Wang, C., Zhang, N., & Wang, C. (2021). Managing Privacy in the Digital Economy. Fundamental Research, 1(5), 543-551.