Integrated Citizen Centered Digital Health and Social Care
A. Värri et al. (Eds.)

doi:10.3233/SHTI200716

Health Data Privacy: Research Fronts, Hot Topics and Future Directions

Javad POOL[a,1], Farhad FATEHI[b,c], Farkhondeh HASSANDOUST[d] and Saeed AKHLAGHPOUR[a]
[a] *The University of Queensland, Brisbane, Australia*
[b] *Monash University, Melbourne, Australia*
[c] *Tehran University of Medical Sciences, Tehran, Iran*
[d] *Auckland University of Technology, Auckland, New Zealand*

Abstract. Health data privacy is an important research stream due to the high impacts on the success of digital health transformation and implementation. Neglecting to safeguard data confidentially and integrity and mitigate risks associated with unauthorized access will lead to failures in materializing benefit from digital health. This study aims to present a bibliometric analysis of health data privacy and provide a platform for future directions. We conducted a literature search between 2010 and 2020 in the Web of Science (WoS) database, resulted in 1,752 records. As part of the bibliometric analysis, concept mapping of health data privacy researches was depicted by network visualization and overlay visualization. These two visualizations represent five research fronts and emerging topics (e.g., digital health, blockchain, the internet of things (IoT)). Finally, we chart directions for future research on health data privacy, highlighting emerging topics, and boundary-breaking alternatives (e.g., GDPR, contact tracing apps in the context of pandemics).

Keywords. Privacy, cybersecurity, digital health, health data, data protection

1. Introduction

Implementation and effective use of digital health can revolutionize healthcare delivery and improve the quality of care. However, unlocking the net benefits of digital health cannot be achieved without protecting the confidentiality and privacy of health data. Investments in health data protection should be included in the healthcare strategy for digital transformations to actualize business value. According to the Cisco data privacy benchmark study in 2020, protecting clients' data drives business value such as innovation, enabling agility and operational efficiency [1]. To gain competitive advantages in the age of Artificial Intelligence (AI), Internet of Medical Things (IoMT), and big data, healthcare industries need to be aligned with updated and new privacy regulations such as the European Union General Data Protection Regulation (GDPR).

During the past decade, the health sector has experienced high profile data breaches. Research highlighted that failures in protecting health data privacy and security are associated with the reputational and financial cost to healthcare organizations and more

[1] Corresponding Author, Javad Pool, The University of Queensland, Brisbane, Australia; E-mail: j.pool@uq.net.au.

importantly, rated to care quality. For example, the Ponemon study reported the average of a health data breach total cost as $6.45 million, which was higher than other industries [2]. Also, a study in the US hospital context revealed that health data breaches were associated with deterioration of care delivery [3].

Health data privacy is closely linked to 'privacy protection practices' and 'security measures'. These two protective safeguards facilitate ensuring the authorized use, confidentiality, and integrity of personal health data. The need for a profound understanding of security and privacy phenomena in the healthcare context has brought together researchers from different domains such as information systems [4] and medical informatics [5]. Research topics in this multidisciplinary field range from social to technical, and psychological perspectives. Therefore, it is important to provide an overview of the published research so that interested academics and practitioners can clearly understand the research profile so far. In this study, therefore, we conducted a bibliometric analysis to examine the academic research fronts in the field of 'Health Data Privacy' to inform scholars and provide impactful directions for future research.

2. Methods

To conduct this bibliometric analysis, first, we developed a search strategy to capture peer-reviewed publications related to 'Health Data Privacy'. Table 1 shows our search query in the Web of Science (WoS) database.

Table 1. Search strategy

Search queries	Limitation
Privacy[a] AND (Health* OR Medic* OR clinic* OR hospital)[b] AND (electronic OR online OR digital OR Internet OR Virtual OR "Information system*" OR "information technolog*" OR "computer*" OR "information and communication technologies" OR ICT)[a]	Years: 2010-2020 Index: SCIE, SSCI, A&HCI, ESCI Type: Articles (excluding reviews) [a] In Title/Abstract/Keywords [b] In Title

The WoS search result was exported in Tab-delimited format as an input for the bibliometric analysis. The analysis has been conducted via VOSViewer version 1.6.15 [6]. Using this software, this study reports a concept mapping via co-occurrences analysis based on authors' keywords (unit of analysis). To demonstrate the meaningful concept mapping, we also created a thesaurus to perform data cleaning. This thesaurus, then, was loaded into VOSViewer to replace or merge synonym terms such as electronic medical records, electronic medical record, electronic medical record (EMR), EMR, and EHR.

3. Results

Our search in the WoS database returned 1,752 records. To provide a concept mapping of health data privacy literature, two types of visualizations, namely 'network visualization' and 'overlay visualization' were represented in our study to illustrate research fronts (privacy-related clusters) and emerging topics.

3.1. Research fronts

Our analysis of co-occurrence of authors' keywords revealed five privacy-related clusters, which are depicted in Figure 1 with different colors.

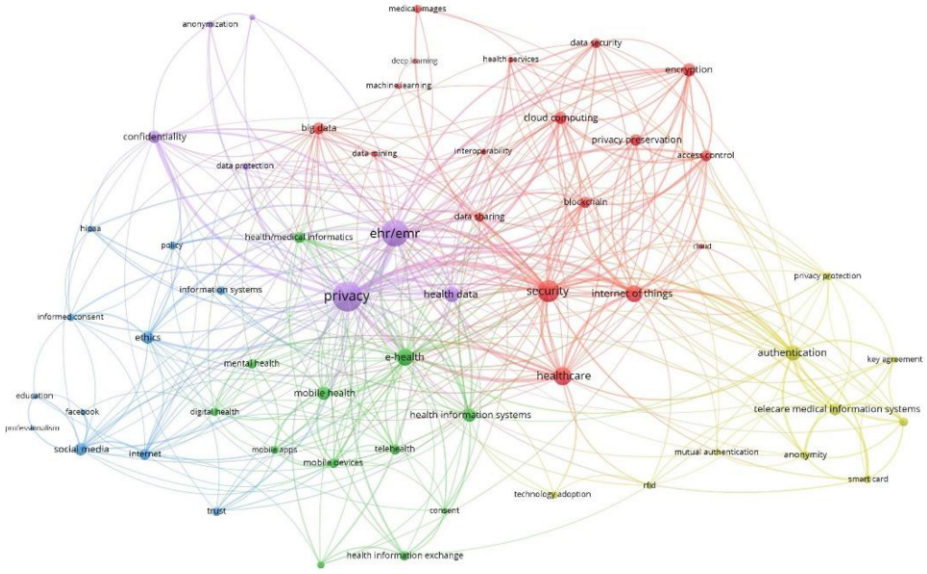

Figure 1. Network visualization of health data privacy. High-res image available at https://bit.ly/2P2rSUk

We labeled these clusters as five research fronts: 'privacy context', 'digital health and care delivery context', 'data right and use aspects', 'security context', and 'technical safeguards and impacts'. Table 2 summarizes these research fronts and hot topics in *health data privacy*.

Table 2. Research fronts and hot topics in health data privacy research

Cluster	Research fronts	Hot topics
#1	Privacy context	Privacy, anonymization, confidentiality, de-identification, EHR/EMR, health data, data protection
#2	Digital health and care delivery context	Digital health, e-health, mobile health, mobile devices, health information exchange, health information systems, mobile apps, medical informatics, mental health, personal health records, consent, telehealth
#3	Data rights and use aspects	Ethics, HIPAA, trust, education, informed consent, policy, professionalism, information systems, internet, social media, Facebook
#4	Security context	Security, big data, data mining, blockchain, cloud, data security, data sharing, deep learning, encryption, health services, healthcare, interoperability, medical images, Internet of Things (IoT), machine learning, privacy preservation, cloud computing
#5	Technical safeguards and impacts	Anonymity, authentication, biometrics, key agreement, mutual authentication, RFID, smart card, telecare medical information systems, privacy protection, technology adoption

As evident in the network visualization, hot topics such as EHR/EMR, and security have received more attention among health data privacy clusters.

3.2. Emerging topics and future directions

Figure 2 illustrates the emerging topics in health data privacy research. The network structure is similar to Figure 1, but the hot topics are colored based on years. The yellow color in the figure indicates hot topics, emerged from 2018 onwards. These 'trending' topics include IoT, blockchain, and digital health. However, topics such as HIPAA, and data mining as a general method are gradually 'cooling off'.

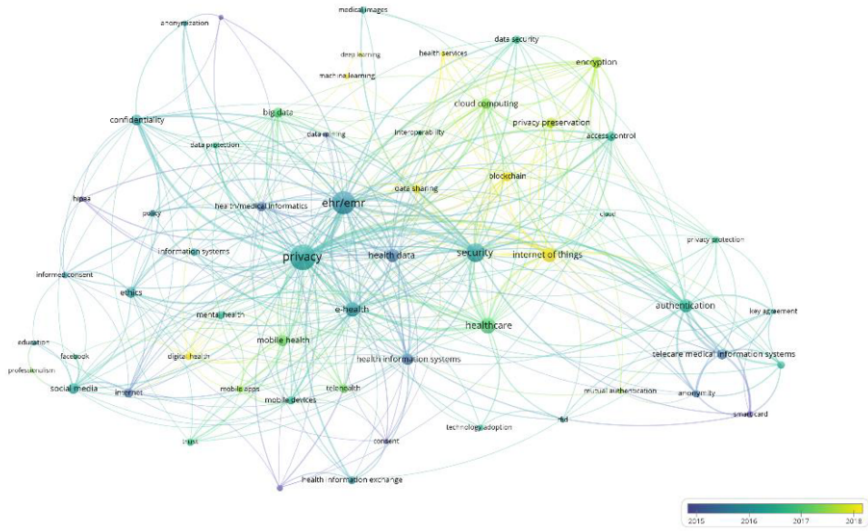

Figure 2. Overlay visualization of health data privacy. High-resolution image is available at https://bit.ly/3jHVg02

Based on the emerging topics (Figure 2) and boundary-breaking alternatives (which can arguably replace the cooling off topics), our suggestions for future research directions are summarized in Table 3.

Table 3. Future research opportunities in health data privacy

	Directions	Key Questions
Emerging Topics	IoT	How and why the introduction and use of IoT in healthcare can increase the risk of health data breaches? How care providers can address and mitigate the cybersecurity and privacy risks associated with the use of IoT in healthcare?
	Blockchain	How do blockchain implementations in healthcare can influence data protection practices? What are the privacy protection opportunities and risks associated with the use of blockchain in healthcare?
	Machine learning and deep learning	How do privacy rights (e.g. the right to restrict processing) can be considered in the use of machine learning and deep learning methods? How do privacy rights (e.g. the right to erasure) affect medical decision making based on these methods?
	Digital health	How can data protection be designed and included in digital health transformations and contribute to improved healthcare performance?
	Data sharing	How do care providers effectively use and implement privacy policies for data sharing in the telehealth context (i.e., GP-to-Specialist, GP-to-Nurse, Patient-to-GP)?

	Directions	Key Questions
		What are the impacts of data breaches on data sharing practices among providers and patients?
	Health services	How can general privacy frameworks (e.g., NIST) be contextually implemented in different health services (e.g., elderly home care)?
Boundary-breaking alternatives	GDPR	How do medical device manufacturers consider GDPR in their processes of designing of digital artefacts? What challenges do Data Protection Officers (DPO) face in protecting health data in practices and how can these data protection challenges be addressed?
	Artificial Intelligence (AI)	How can AI play a role in detecting unauthorized access to health data and facilitate response to health data breaches?
	5G Internet	How do privacy concerns related to health data will influence the adoption of 5G Internet in healthcare?
	Privacy protection value	How can healthcare providers plan and actualize business value from protecting patient data (e.g., innovation, agility)?
	Contact tracing apps	How can 'privacy concerns' and 'lack of data protection by design' trigger individuals' resistance to adopt and use contact tracing apps in the pandemic context such as COVID-19?

4. Conclusions

Healthcare organizations and users are moving towards digital health to enhance their performance and co-create value (e.g., in improved management of chronic diseases). However, unlocking the net benefit of digital health technologies requires attention to and practice of safeguarding health data and protecting users' privacy. This bibliometric study reported on the hot topics in health data privacy literature. Furthermore, our study illustrated that the research streams are moving towards new trends such as blockchain and IoT, which show opportunities for health data privacy researches. Also, beyond the emerging trends, we proposed new directions for privacy researches, i.e., GDPR and privacy in the context of contact tracing apps in pandemics. Theses research opportunities are worth exploring to inform research, policy, and data protection practices. While future research can update the hot research topics in the five identified clusters of health data privacy, we encourage scholars to set a high priority for emerging topics and delve deeper into boundary-breaking alternatives.

References

[1] CISCO. From Privacy to Profit: Achieving Positive Returns on Privacy Investments. USA: Cisco and/or its affiliates, 2020.
[2] Ponemon-Institute. Cost of a Data Breach. USA: Ponemon Institute LLC, Results sponsored, analyzed and reported by IBM Security, 2019.
[3] Choi SJ, Johnson ME, Lehmann CU. Data breach remediation efforts and their implications for hospital quality. Health services research. 2019;54(5):971-80.
[4] Kim SH, Kwon J. How Do EHRs and a Meaningful Use Initiative Affect Breaches of Patient Information? Information Systems Research. 2019;30(4):1184-202.
[5] Prasser F, Spengler H, Bild R, Eicher J, Kuhn KA. Privacy-enhancing ETL-processes for biomedical data. International journal of medical informatics. 2019;126:72-81.
[6] Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. scientometrics. 2010;84(2):523-38.