Data Privacy Issues and Challenges: Lessons from Literature

Renu Bala*
Research Scholar, CDLU, Sirsa, Haryana, India
*Email: renu820kamboj@gmail.com

ABSTRACT

Through this paper, we strived to find out the challenges and ethical issues in the stream of data privacy in the present era of the twenty-first century, based on old studies. To review the literature, a range of sources were searched on the topic and fifty-one research papers have been taken as a sample, out of which nine are from USA and India each and eight form Philippines. The present study focuses on empirical studies which were published during the period 2008-2020. Different kinds of journals, magazines, articles, newspaper, websites, etc. have been explored to analyze ancient data for research purpose. During the research, efforts have been made to identify the challenges as well as ethical issues which the researchers have to face during the research process. In the research, it has been founded that there exist a lot of challenges for the data privacy and it becomes very difficult to follow the ethics to maintain the privacy and security of data in this technology-driven era. In the end, based on available data it has been concluded that there is a great need of hard punishments for those people who misuse the data for their own sake and demoralize those people who make hard efforts to carry out the research and generate new concepts and ideas. So the government should take necessary measures toward data privacy area for security and safety purpose.

Keywords: Privacy, Data Sharing, Protection, Ethics, Challenges.

INTRODUCTION

Privacy, security and trust these three terms are interlinked with each other in such a way how the law and ethics are related to each other. The term data privacy represents how the data should be collect, use and access by considering the legal rights (Lee *et al*., 2016). On the other side, ethics denotes those responsibilities, sometimes which become the obligation to pursue (Knoppers & Thorogood, 2017). At present, in this technology-driven era internet-based research is very popular and most of the researchers follow this method widely to fulfil the purpose of data collection and analysis. The umbrella of Internet-based research cover video conferencing for interviews, online surveys, analysis of 'e-conversations', web page content analysis, discussion blogs, chat rooms and email etc. (Cox, 2012). With the passage of the time, a large number of social and technologies changes can be noticed and these changes are altering the scale and scope of the data for researchers (Fiesler, 2019). Instead of a lot of advantages of the latest technology, some drawbacks also exist. Data sharing and data storage provide different kinds of the benefits to researchers but it creates a lot of hurdles also from the security perspective of data and in this present scenario, traditional privacy mechanism and security system are not adequately sufficient to cope with data explosion (Venkatraman & Venkatraman, 2019). As a result, researchers have to face different kinds of data issues during their authentic work such as modification of secondary data, storing of data, sharing of data etc. and in these situations ethical decision making take place (Boyd *et al*., 2016). Due to the advanced technology, the chances of misuse of data are increasing day by day and due to this situation, it becomes very hard to follow the ethics (Hand, 2018). Researchers have to face various types of data privacy challenges during their work such as securing the computations under distributed programming frameworks, securing data storage and the transactional logs, data provenance, end-point filtration and validation, real-time secure monitoring (Mehta & Rao 2015) etc. So

the government and other private institutions should focus to develop different kinds of new software under the latest technology, which can protect data by maintaining its privacy and avoiding misuse, modification and wrong interpretation etc.

RESEARCH OBJECTIVES

The main objectives of the study are to identify and examine the challenges in the stream of data privacy and ethical dilemma in this present era of the 21st century. To achieve the purpose of the study efforts have been made through the review of published research work in the same area.

DATA AND METHODOLOGY

Data

This research subsumes the review of fifty-one research papers on the challenges of data privacy and ethical dilemma, published during the years 2008-2020. Due to a short period, all the studies in the respective field could not be covered.

Methodology

In this paper, efforts have been made to present the dynamics of research work published on challenges in the field of data privacy and ethical dilemma. The keywords like "privacy, challenges, ethics and dilemma" etc. have been used for literature search. In the earlier studies, different kinds of statistical tools like t-test, ANOVA, factor analysis, correlation and regression analysis, percentage analysis, etc. have been used for the analysis of data. For the research purpose, hundreds of papers were meticulously studied and finally reduced to fifty-one research papers to carry out the present study. In the end, data have been analyzed with the help of tables to draw results and valid conclusions.

CLASSIFICATION OF LITERATURE

The reviewed literature has been divided into different kinds of categories based on various parameters such as year, country as well as source used for data collection as elaborated below.

a)   Year based classification

b)   Country-based classification

c)   Source-based classification

Year based classification

The frequency of the research work conducted in several years (Table 1) presents that the research in the similar stream gained popularity after the year 2011, because a lot of research work has been conducted in the same area during the year 2012, and then in 2014 and afterwards.

Table 1. Year-wise Classification.

S. No.	Year(s)	Frequency	S. No.	Year(s)	Frequency
1	2008	01	8	2015	03
2	2009	01	9	2016	07
3	2010	01	10	2017	05
4	2011	01	11	2018	05
5	2012	05	12	2019	10
6	2013	02	13	2020	02
7	2014	08	14		
Total					**51**

Source. Compiled from earlier studies.

Country-based classification

The country-based classification of the study (Table 2) exhibits that most of the researches have been conducted in the USA, India and the Philippines. Out of fifty-one research papers, nine papers are from USA and India each, eight papers are from the Philippines, and five papers belong to the UK. Instead of these countries, some other countries such as Europe, Africa and New York etc. have also presented their interest in the respective field.

Table 2. Country-wise Classification.

S. No.	Country Name	No. of Studies	S. No.	Country Name	No. of Studies
1	USA	09	9	Malaysia	02
2	India	09	10	Hong Kong	01
3	Philippines	08	11	Canada	01
4	UK	05	12	China	01
5	Europe	04	13	Saudi Arabia	01
6	Africa	03	14	Singapore	01
7	New York	03	15	Switzerland	01
8	Australia	02			
Total					**51**

Source-based classification

Table 3 presents the name of several journals, which have been used to fulfil the purpose of data collection as well as the name of some papers, which have been presented in an international conference. Based on available data it has been inferred that four papers have been selected from the Journal 'Advancing Information Technology Research', three papers have been chosen from the journal 'Asia Pacific Journal of Multidisciplinary Research' and two papers have been taken from the 'BENCHMARK: College of Business Research Journal' and 'Journal of Business Research' each and so on to collect the required data.

Table 3. Source-wise Classification

Particulars	No. of Papers
(A) Journals	
Advancing Information Technology Research	04
Asia Pacific Journal of Multidisciplinary Research	03
BENCHMARK: College of Business Research Journal	02
Journal of Business Research	02
IEEE	01
WILEY	01
PLOS ONE	01
Big Data	01
ISACA Journal	01
AIMS Mathematics	01
Asian Bioethics Review	01
Health Services Research	01
Current Opinion in Systems Biology	01
Procedia Computer Science	01
Data & Society	01

Internet Policy Review	01
Frontiers in Big Data	01
Systemics, Cybernetics and Informatics	01
Journal of Big Data	01
Journal of Law and the Biosciences	01
Journal of Medical Internet Research	01
Computer Communication Review	01
Practitioner Research in Higher Education	01
Proceedings of the ACM on Human-Computer Interaction	01
Qualitative Health Research	01
Mara International Journal of Scientific & Research Publications	01
Philosophical Transactions of the Royal Society	01
International Data Privacy Law	01
International Journal of Computer Trends and Technology	01
International Journal of Computer Sciences and Engineering	01
International Journal of Engineering Science and Technology	01
International Journal of Control Theory and Applications	01
International Journal of Network Security & its Applications (IJNSA)	01
International Journal of Computer Engineering & Technology (IJCET)	01
International Journal of Engineering Development and Research	01
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)	01
International Journal of Innovative Research in Advanced Engineering (IJIRAE)	01
International Journal of Computer Science and Information Technology & Security (IJCSITS)	01
IFLA (International Federation of Library Associations and Institutions) Journal	01
International Journal of Innovative Technology and Exploring Engineering (IJITEE)	01
(B) International Conference Paper & website	
Paper presented at First USENIX Workshop on Large-Scale Exploits and Emergent Threats, San Francisco, CA, USA, Proceedings	01
Paper presented at Workshop on Secure Data Management	01
Available at https://www.researchgate.net	02
TOTAL	**51**

Source. Compiled from earlier studies.

REVIEW AND ANALYSIS

This section shows the findings of the literature reviewed, to fulfil the objectives of the respective study.

Primary Focus

The primary focus of the present study is to identify as well as examine the challenges in the stream of data privacy in the present era of the 21st century. Different authors like Burstein (2008), Padhy *et al*. (2011), Bertino (2014), Mehta & Rao (2015), Bhandari *et al*. (2016), Shomrani *et al*. (2018), Chiauzzi & Wicks (2019) etc. focused on the challenges in data privacy (Table 4) as presented below.

Table 4. Challenges in the Stream of Data Privacy.

S. No.	Challenges in data privacy	S. No.	Challenges in data privacy
1.	Share and publish network traces	11.	Data provenance
2.	Collection of network traces	12.	End-point filtration and validation
3.	Data encryption	13.	Real-time secure monitoring
4.	Access controls	14.	Input validating and filtration
5.	Server consolidation	15.	Granular accessed control
6.	Protection from the insider threats	16.	Unsecured data storage
7.	Data trustworthiness	17.	Cryptographically enforce data-centric security
8.	Reconcile data security and privacy	18.	Scalable privacy-preserving analytics and data mining
9.	Securing computed data under the distributed programming frameworks	19.	Repetition of scraping by multiple accounts
10.	Securing data storage and the transactional logs	20.	Researcher misrepresentation as well as fake profile data

Source. Compiled from earlier studies.

Table 4 exhibits that there exist different kinds of challenges in the data privacy and these highlighted challenges are share and publish network traces, collection of network traces (*Burstein, 2008*), data encryption, access controls, server consolidation (Padhy *et al*., 2011), protection from the insider threats, data trustworthiness, reconcile data security and privacy (Bertino, 2014), securing the computations under distributed programming frameworks, securing data storage and the transactional logs, data provenance, end-point filtration and validation, real-time secure monitoring (Mehta & Rao 2015), input validating and filtration, granular accessed control, unsecured data storage (Bhandari *et al*., 2016), cryptographically enforce data-centric security, scalable privacy-preserving analytics and data mining (Shomrani *et al*., 2018), repetition of scraping by multiple accounts, researcher misrepresentation as well as fake profile data (Chiauzzi & Wicks, 2019) etc.

Secondary Focus

The secondary focus of the present research is to identify the ethical issues in data privacy in the 21st century. Different types of research papers, articles, and journals etc. have been reviewed to search out the ethical issues in data privacy. These ethical issues can be noticed in different kinds of forms (Table 5).

Table 5. Ethical Issues in Data Privacy in the 21st Century.

S. No.	Ethical issues	S. No.	Ethical issues
1.	Misuse of social licensing	6.	Negative exploitation of research tools, techniques, and standards *etc.*
2.	Make a balance between risks and benefits	7.	Misuse of data
3.	Wrong use of 'right to science' in international law	8.	Break the privacy breach
4.	Wrong governance of big data	9.	Wrong interpretation of data as per requirements
5.	Modification of cybersecurity research ethics	10.	Consider the research as a source of money only

Source. Compiled from earlier studies.

Table 5 presents that there exist several types of ethical issues in the data privacy and the major issues include misuse of social licensing (Xafis *et al*., 2019), make the balance between risks and benefits, wrong use of 'right to science' in the international law, wrong governance of big data (Knoppers & Thorogood, 2017), modification of cyber security research ethics, negative exploitation of research tools, techniques, and standards etc. (Kenneally & Bailey, 2014), misuse of data, break the privacy breach (Hand, 2018), wrong interpretation of data as per requirements, consider the research as a source of money only (Boyd *et al*., 2016) etc.

Other Observations and Findings

During the research, it has been observed that the number of studies have been conducted in this stream and the main focused on such issues has been paid after the year 2011. In this area, various researches and surveys etc. have been conducted at the international level due to its importance as per the present scenario. It has also been founded that different kinds of challenges have to be faced by the researchers to keep secure their data and maintain privacy. Various ethical issues also exist in the data privacy which plays a very important role to secure the data at every level of work. To maintain the data security various tools and techniques are introduced by the organizations from time to time. The latest technology is playing a very significant role to maintain and secure the data at different levels.

CONCLUSION

In the current study, fifty-one research papers form the year 2008 to 2020 were collected and deeply analyzed. A lot of websites, online databases, journals, and websites etc. have been reviewed to identify the challenges and ethical issues in the data privacy field. After the analyses of collected data, it has been revealed that the existing challenges in the data security ruin the actual meaning of data and lost its authenticity. Researchers, who make serious and hard efforts in their work, feel demoralized due to theft and misuse of actual data. Actually data privacy has become a major issue in this era and to fight against such types of issues, private and government organizations are paying attention and making good efforts to introduce new kinds of techniques and software etc. for data security. But strict rules and punishment should be fixed for those people who exploit the original work and misuse the actual data.

The Utility of the Study

This study presents an organized and detailed review of empirical studies in the field of data privacy challenges and ethical issues. It will provide a reference to the researchers to carry out their research in this respective field and provide a clear picture of data privacy challenges. After analyzing the above-mentioned challenges and ethical issues, this study will also provide knowledge to the researchers regarding the present challenges

of the respective fields. After the analysis of difficulties arisen in this area, solutions can be found out to solve the problems and overcome the challenges.

REFERENCES

1. Abacahin, J. & Amil, D. *et al*. (2012). Sales and Inventory System of Edmar Marketing. Advancing Information Technology Research, 3(1), 259-282, ISSN: 2094-9626, available at https://docs.google.com.
2. Abouelmehdi, Karim, Abderrahim Beni-Hessane, and Hayat Khaloufi. "Big healthcare data: preserving security and privacy." Journal of Big Data 5.1 (2018): 1-18.
3. Agapay, Riel S., *et al*. "Online Book Information System of Liceo de Cagayan University." Advancing Information Technology Research 3.1 (2012): 1-1.
4. Ahmad, S. & Azam, M. *et al*. (2018). A Survey on Security and Privacy of Big Data. International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), 5(4), 96-100, ISSN: 2394-2320, available at https://www.researchgate.net.
5. Alnajrani, Hussain Mutlaq, Azah Anir Norman, and Babiker Hussien Ahmed. "Privacy and data protection in mobile cloud computing: A systematic mapping study." PloS one 15.6 (2020): e0234312.
6. Andersen, Anders, and Merete Saus. "Privacy preserving distributed computation of community health research data." Procedia computer science 113 (2017): 633-640.
7. Rado, Marco Marvin, and Brain D. Angeles. "LAN-Based Management Information System of Petron Gas Service Station-Puerto Branch." Advancing Information Technology Research 3.1 (2012): 1-1.
8. Bao, R. & Chen, Z. *et al*. (2017). Challenges and Techniques in Big Data Security and Privacy: A Review. WILEY, 1(4), 1-8, ISSN: 2048-416X, available at https://onlinelibrary.wiley.com.
9. Bertino, E. (2014). Data Security – Challenges and Research Opportunities. Paper presented at Workshop on Secure Data Management, 1-5, available at https://www.cs.purdue.edu.
10. Bhandari, R. & Hans, V. *et al*. (2016). Big Data Security – Challenges and Recommendations. International Journal of Computer Sciences and Engineering, 4(1), 93-98, ISSN: 2347-2693, available at https://www.researchgate.net.
11. Burstein, A. J. (2008). Conducting Cyber Security Research Legally and Ethically. Paper presented at First USENIX Workshop on Large-Scale Exploits and Emergent Threats, San Francisco, CA, USA, Proceedings. Available at https://www.usenix.org.
12. Buted, Dexter R., *et al*. "Effects of social media in the tourism industry of Batangas Province." Asia Pacific Journal of Multidisciplinary Research| Vol 2.3 (2014): 123-131.
13. Boyd, Danah, EMILY F. Keller, and B. O. N. N. I. E. Tijerina. "Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research." Data & Society (2016).
14. Chiauzzi, Emil, and Paul Wicks. "Digital trespass: ethical and terms-of-use violations by researchers accessing data from an online patient community." Journal of medical Internet research 21.2 (2019): e11985.
15. Cooke, Louise. "Privacy, libraries and the era of big data." (2018): 167-169.
16. Fabiano, N. (2019). Ethics and the Protection of Personal Data. Systemics, Cybernetics and Informatics, 17(2), 58-64, ISSN: 1690-4524, available at https://www.researchgate.net.
17. Floridi, L. & Taddeo, M. (2016). What is Data Ethics? Philosophical Transactions of the Royal Society, 374(2083), 1-8, ISSN: 2053-9223, available at https://www.researchgate.net.
18. Hand, David J. "Aspects of data ethics in a changing world: Where are we now?" Big data 6.3 (2018): 176-190.
19. Hasselbalch, Gry. "Making sense of data ethics. The powers behind the data ethics debate in European policymaking." Internet Policy Review 8.2 (2019): 1-19.
20. Inukollu, Venkata Narasimha, Sailaja Arsi, and Srinivasa Rao Ravuri. "Security issues associated with big data in cloud computing." International Journal of Network Security & Its Applications 6.3 (2014): 45.
21. Kaiser, Karen. "Protecting respondent confidentiality in qualitative research." Qualitative health research 19.11 (2009): 1632-1641.
22. Kantarcioglu, Murat, and Elena Ferrari. "Research challenges at the intersection of big data, security and privacy." Frontiers in Big Data 2 (2019): 1.
23. Kante, M. (2017). A Review of Big Data Security and Privacy Issues. Mara International Journal of Scientific & Research Publications, 1(1), 49-54, ISSN: 2523-1456, available at https://www.researchgate.net.
24. Kenneally, Erin, and Michael Bailey. "Cyber-security research ethics dialogue & strategy workshop." ACM SIGCOMM Computer Communication Review 44.2 (2014): 76-79.

25. Knoppers, B. M. & Thorogood, A. M. (2017). Ethics and Big Data in Health. Current Opinion in Systems Biology, 4(1), 53-57, ISSN: 24523100, available at https://www.sciencedirect.com.

26. Lane, J. & Schur, C. (2020). Balancing Access to Health Data and Privacy: A Review of the Issues and Approaches for the Future. Health Services Research, 45(5), 1456-1467, ISSN: 00179124, available at https://www.ncbi.nlm.nih.gov.

27. LEE, Wanbil W., Wolfgang ZANKL, and Henry CHANG. "An ethical approach to data privacy protection." Isaca Journal (2016).

28. Magno, K., and J. Pabico. "Digital Anthropometry: Model, Implementation, and Application." Asia Pacific Journal of Multidisciplinary Research 2.3 (2014).

29. Mehmood, Abid, *et al*. "Protection of big data privacy." IEEE access 4 (2016): 1821-1834.

30. Mehta, B. & Rao, U. P. (2015). Big Data Privacy: Issues and Challenges. Project: Privacy Preserving Big Data Publishing, available at https://www.researchgate.net.

31. Mikalef, Patrick, *et al*. "Big data analytics and firm performance: Findings from a mixed-method approach." Journal of Business Research 98 (2019): 261-276.

32. Moura, José, and Carlos Serrão. "Security and privacy issues of big data." Handbook of research on trends and future directions in big data and web intelligence. IGI Global, 2015. 20-52.

33. Nuyda, Darryl M., *et al*. "Online Inventory and Monitoring System of Cagayan de Oro City Health Office." Advancing Information Technology Research 3.1 (2012): 1-1.

34. Padhy, Rabi Prasad, Manas Ranjan Patra, and Suresh Chandra Satapathy. "Cloud computing: security issues and research challenges." International Journal of Computer Science and Information Technology & Security (IJCSITS) 1.2 (2011): 136-146.

35. Padmapriya, S. & Partheeban, N. *et al*. (2019). Enhanced Cyber Security for Big Data Challenges. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(10), 3478-3481, ISSN: 2278-3075, available at https://www.ijitee.org.

36. Rad, B. B. & Akbarzadeh, N. *et al*. (2016). Security and Privacy Challenges in Big Data Era. International Journal of Control Theory and Applications, 9(43), 437-448, ISSN: 09745572, available at https://www.researchgate.net.

37. Ravi, Renjith V., and R. Dr Mahalakshmi. "Universal Intelligent Data Encryption Standards: A Review." Asia Pacific Journal of Multidisciplinary Research 2.3 (2014): 1-1.

38. Resare, Princess Mira M., and Mark Anthony P. Bautista. "STUDENTS'RECORD INFORMATION SYSTEM (SRIS) AIMS-CeSS Department." BENCHMARK: College of Business Research Journal 2.1 (2014): 1-1.

39. Rubinstein, I. S. (2013). Big Data: The End of Privacy or a New Beginning? International Data Privacy Law, 3(2), 74-87, ISSN: 2044-4001, available at https://watermark.silverchair.com.

40. Scheibner, James, *et al*. "Data protection and ethics requirements for multisite research with health data: a comparative examination of legislative governance frameworks and the role of data protection technologies." Journal of Law and the Biosciences 7.1 (2020): lsaa010.

41. Xafis, Vicki, *et al*. "An ethics framework for big data in health and research." Asian Bioethics Review 11.3 (2019): 227-254.