



(RESEARCH ARTICLE)



Data and science engineering: The ethical dilemma of our time-exploring privacy breaches, algorithmic biases, and the need for transparency

Shubham Shubham ¹, Saloni Saloni ² and Sidra-Tul-Muntaha ^{3,*}

¹ *Israel Institute of Technology and a master's degree in building construction from Georgia Institute of Technology, Georgia.*

² *KU Leuven in Belgium, Belgium.*

³ *Fatima Jinnah Women University and also graduated from the University of People America. America.*

World Journal of Advanced Research and Reviews, 2023, 18(01), 762–768

Publication history: Received on 08 March 2023; revised on 15 April 2023; accepted on 18 April 2023

Article DOI: <https://doi.org/10.30574/wjarr.2023.18.1.0677>

Abstract

This paper explores the ethical dilemmas associated with data and science engineering, with a focus on privacy breaches, algorithmic biases, and the need for transparency. With the increasing reliance on data-driven decision making and machine learning algorithms, the ethical implications of these technologies have become a pressing issue in various sectors. The study aimed to identify the most significant ethical concerns, analyze their impact on society, and provide solutions to address these issues.

The research utilized a systematic review of 18 studies to identify the key ethical issues in data and science engineering. The findings revealed that privacy breaches, algorithmic biases, and lack of transparency were the most prevalent ethical concerns. These issues can have significant implications for individuals and groups, including discrimination, loss of autonomy, and reputational harm. The study also identified vulnerable groups, such as marginalized communities, who may be disproportionately affected by these issues.

To address these ethical concerns, the study proposed several solutions, including the development of ethical guidelines, increased transparency and accountability, and the use of diverse and representative datasets. The solutions were informed by the literature review, case studies, and analysis of real-world examples. The study also assessed the feasibility of implementing these solutions and highlighted potential barriers to implementation.

Keywords: Data; Science engineering; Ethics; Privacy; Transparency; Bias; Algorithms

1. Introduction

With organisations depending on data-driven decision-making and machine learning algorithms to obtain a competitive edge, the usage of data and scientific engineering has been growing quickly in recent years. Although these technologies have numerous advantages, they also bring up ethical issues that need to be resolved. The gathering, processing, and analysis of data in various sectors can give rise to ethical problems including bias, privacy, and transparency.

1.1. Ethics

Ethics, in general, refers to a set of moral principles and values that guide human behavior and decision-making. Ethics plays a crucial role in various aspects of human life, including personal and professional domains. In the context of data and science engineering, ethics becomes a critical consideration in the collection, processing, and analysis of data [1]. The field of data and science engineering involves the use of various technologies to collect, store, process, analyze, and

* Corresponding author: Sidra-Tul-Muntaha

visualize data. The ethical implications of data and science engineering arise from the fact that the data being used often belongs to individuals or groups, and the decisions made based on that data can have significant impacts on their lives [2].

1.2. Ethics in Data and Science Engineering

Ethics in the field of data and science engineering refers to the moral principles and values that guide the responsible and ethical use of data and technology. It involves a range of issues, including privacy, transparency, bias, and accountability. The use of data and science engineering has the potential to impact people's lives in many ways, from the way businesses operate to the way individuals are treated by governments or other organizations [3]. Thus, it is essential to consider the ethical implications of data and science engineering and ensure that these technologies are used responsibly and with respect for individuals' rights and interests.

1.3. Importance of Ethics in Data and Science Engineering

The importance of ethics in data and science engineering cannot be overstated. These technologies have the potential to improve our lives in many ways, from healthcare to education to transportation. However, they also raise significant ethical concerns, particularly in terms of privacy, transparency, and bias. The responsible and ethical use of data and science engineering can help ensure that the benefits of these technologies are realized without causing harm or violating individuals' rights [4]. Additionally, ethical considerations can help build trust and credibility in these technologies, ensuring that they are used effectively and sustainably over the long term. Overall, the importance of ethics in data and science engineering lies in its ability to balance the potential benefits of these technologies with the need to ensure that they are used ethically and responsibly [5].

The purpose of this study is to investigate how data and science engineering ethical concerns affect society. The main aims of this study are to identify the most important ethical problems in data and science engineering, analyse how they affect society, and provide remedies.

1.4. Research Objectives

The research objectives of this study are as follows:

- To identify the key ethical issues in data and science engineering: This objective will involve conducting a thorough literature review and analyzing existing research to identify the ethical issues that arise from the use of data and science engineering.
- *To examine the impact of these ethical issues on society:* This objective will involve conducting case studies and analyzing real-world examples to understand how these ethical issues affect society as a whole and to identify groups that are particularly vulnerable to these issues.
- *To propose solutions to address these ethical issues:* This objective will involve developing recommendations and guidelines for addressing the ethical issues identified in the previous objectives. These solutions will be informed by the literature review, case studies, and analysis of real-world examples.
- *To assess the feasibility of implementing these solutions:* This objective will involve assessing the practicality and feasibility of implementing the proposed solutions in real-world contexts. This will involve analyzing the potential costs and benefits of these solutions, as well as potential barriers to implementation.

1.5. Research Questions

To achieve these objectives, this research will address the following research questions:

- What are the key ethical issues in data and science engineering?
- How do these ethical issues affect society as a whole, and are certain groups more affected than others?
- What solutions can be proposed to address these ethical issues, and how can they be implemented effectively?

By answering these research questions, this study will provide valuable insights into the ethical challenges that arise from the use of data and science engineering. It will also offer recommendations for how organizations and policymakers can address these challenges to ensure that these technologies are used ethically and responsibly.

1.6. Significance of Study

This study is significant for several reasons. Firstly, it will contribute to the ongoing conversation about the ethical implications of data and science engineering. As these technologies become more prevalent, it is essential that we

understand the ethical challenges they pose and work to address them. Secondly, this study will provide valuable insights into the impact of these ethical issues on society. By examining real-world examples and case studies, we can better understand how these issues affect different groups and develop more targeted solutions to address them. Finally, this study will offer practical recommendations for addressing these ethical issues. By assessing the feasibility of implementing these solutions, we can ensure that they are realistic and effective in addressing the challenges posed by data and science engineering.

1.7. Limitations of Study

This study also has several limitations that should be acknowledged. Firstly, the research will be limited by the availability and quality of existing literature on the topic. While efforts will be made to conduct a thorough literature review, it is possible that some relevant studies may be missed. Secondly, the research may be limited by the availability of real-world examples and case studies. While efforts will be made to identify and analyze as many examples as possible, it is possible that some groups or issues may be underrepresented in the available data.

Finally, the research may be limited by the practicality of implementing the proposed solutions. While efforts will be made to assess the feasibility of these solutions, it is possible that they may not be practical or effective in all contexts. Despite these limitations, this study will provide a valuable contribution to the ongoing conversation about the ethical implications of data and science engineering, and will offer practical recommendations for addressing these challenges.

2. Literature Review

2.1. Introduction

The literature review aims to provide an overview of the existing literature on the ethical issues surrounding data and science engineering. The chapter begins by defining data and science engineering, followed by a discussion of the ethical implications of these technologies. The literature review also explores the different ethical issues that arise from the use of data and science engineering, including privacy breaches, algorithmic biases, and the need for transparency. Finally, the chapter concludes with a discussion of the current state of research on these topics and gaps in the existing literature.

2.2. Data and Science Engineering

Data and science engineering is a field that involves the use of technology to collect, store, process, analyze, and visualize data. This field encompasses a wide range of technologies, including machine learning algorithms, data analytics tools, and data management systems. The use of data and science engineering has become increasingly prevalent in various sectors, including healthcare, finance, and marketing, among others [6].

2.3. Ethical Implications of Data and Science Engineering

The use of data and science engineering has given rise to several ethical concerns. These concerns arise from the fact that data often belongs to individuals or groups, and the decisions made based on that data can have significant impacts on their lives [7]. Therefore, it is essential to consider the ethical implications of data and science engineering and ensure that these technologies are used responsibly and with respect for individuals' rights and interests.

2.4. Privacy Breaches

One of the most significant ethical issues surrounding data and science engineering is privacy breaches. Privacy breaches occur when data that is supposed to be kept private is exposed or made public without the individual's consent [8]. This can happen due to data breaches, hacking, or unauthorized access. Privacy breaches can have severe consequences, including identity theft, financial loss, and reputational damage [9].

2.5. Algorithmic Biases

Algorithmic biases are another ethical issue that arises from the use of data and science engineering [10]. Algorithmic biases occur when machine learning algorithms make decisions based on biased data sets. This can lead to unfair treatment of individuals or groups based on their race, gender, or other characteristics [11]. Algorithmic biases can have significant consequences, including perpetuating social inequality and discrimination [12].

2.6. Need for Transparency

The need for transparency is another ethical issue that arises from the use of data and science engineering. Transparency refers to the requirement that organizations and individuals who use data and science engineering technologies should be open and honest about their data collection and use practices [13]. Lack of transparency can lead to distrust and suspicion, particularly in cases where data is used for decision-making that affects individuals or groups [14].

2.7. Impact of Ethical Issues on Society

The ethical issues in data and science engineering have significant impacts on society as a whole. These impacts include, loss of Trust Privacy breaches, algorithmic bias, and the lack of transparency can all lead to a loss of trust in data and science engineering technologies. This can make it more challenging for organizations to use these technologies effectively, as individuals may be reluctant to share their data or participate in data-driven decision-making processes [15].

2.7.1. Reinforcement of Social Injustice

Algorithmic bias can lead to the reinforcement of social injustice by perpetuating existing biases and discriminating against certain groups of people. This can result in individuals being treated unfairly in various contexts, including employment, healthcare, and criminal justice [16].

2.7.2. Violation of Individual Rights

Privacy breaches can result in individuals' sensitive information being exposed, which can violate their right to privacy. The lack of transparency can also violate individuals' rights by making it difficult for them to understand and contest decisions made using algorithms [17].

2.8. Current State of Research

The existing literature on the ethical issues surrounding data and science engineering is still in its infancy. However, researchers are increasingly recognizing the need to address these ethical concerns to ensure that these technologies are used responsibly and with respect for individuals' rights and interests. Much of the current research has focused on specific ethical issues, such as privacy breaches or algorithmic biases, rather than examining the broader ethical implications of data and science engineering [18].

2.9. Gaps in the Literature

Despite the growing interest in the ethical implications of data and science engineering, there are still several gaps in the existing literature. One significant gap is the lack of research on the ethical challenges that arise from the use of data and science engineering in developing countries. Another gap is the need for more research on the practical implementation of ethical guidelines and recommendations in real-world contexts.

2.10. Conclusion

In conclusion, the literature review has highlighted the importance of ethics in data and science engineering and the ethical concerns that arise from the use of these technologies. The review has identified privacy breaches, algorithmic biases, and the need for transparency as the most significant ethical challenges in data and science engineering. The review has also shown that these challenges can have significant impacts on individuals and society as a whole, particularly vulnerable groups.

3. Material and method

This study is based on a review of 18 previous studies that explore the ethical concerns of data and science engineering, including privacy breaches, algorithmic biases, and the need for transparency. The aim of this research is to identify the key ethical issues in data and science engineering, understand their impact on society, and propose solutions to address them.

3.1. Research Design

The research design for this study is a systematic literature review. The systematic review is a well-established method for synthesizing research evidence on a particular topic in a rigorous and reproducible manner. This design allows for

a comprehensive and unbiased review of existing research, enabling researchers to identify key themes and trends in the literature.

3.2. Search Strategy

The search strategy for this study involved a systematic search of four academic databases: Scopus, Web of Science, PubMed, and IEEE Explore. The search was conducted in April 2023 using the following keywords: "data ethics," "algorithmic bias," "privacy breaches," and "transparency." Boolean operators "AND" and "OR" were used to combine the keywords. The search was limited to studies published in English between 2010 and 2023.

3.3. Inclusion and Exclusion Criteria

The inclusion criteria for this study were studies that focus on the ethical concerns of data and science engineering, including privacy breaches, algorithmic biases, and the need for transparency. Only studies published in English between 2010 and 2023 were included. Studies that were not peer-reviewed, did not focus on the ethical concerns of data and science engineering, or were not available in full text were excluded from this review.

Overall, the systematic review methodology employed in this study provides a robust and reliable approach for synthesizing existing research on the ethical concerns of data and science engineering. The inclusion and exclusion criteria ensure that only high-quality, relevant studies are included in the review, and the thematic analysis approach enables the identification of key themes and trends in the literature.

4. Result and Discussion

This study aimed to investigate the ethical concerns arising from the use of data and science engineering and their impact on society. Through a review of 18 previous studies, we identified three main ethical issues: privacy breaches, algorithmic biases, and the need for transparency. Researchers also found that these issues have significant impacts on individuals and society, particularly vulnerable groups such as minorities and low-income communities.

Regarding privacy breaches, we found that the collection, processing, and storage of personal data can put individuals at risk of identity theft, financial fraud, and other forms of exploitation. In some cases, these breaches have resulted in significant harm to individuals, including reputational damage and emotional distress. To address this issue, researchers recommend that organizations be transparent about their data collection practices and obtain explicit consent from individuals before collecting their data. Additionally, data should be securely stored and processed to prevent unauthorized access and use.

Algorithmic biases were another significant ethical concern identified in this study. Researchers found that algorithms can perpetuate and amplify biases in society, particularly regarding race, gender, and socioeconomic status. These biases can lead to unfair treatment and discrimination in areas such as hiring, lending, and law enforcement. To address this issue, researchers recommend that organizations assess and monitor their algorithms for bias, using diverse and representative datasets to ensure fairness and accuracy. Additionally, algorithmic decision-making should be subject to human oversight and review to mitigate the risk of harmful biases.

Finally, researchers found that the need for transparency is a critical ethical concern in data and science engineering. Lack of transparency can lead to mistrust and suspicion of organizations using these technologies, particularly in cases where algorithms are used to make significant decisions affecting individuals' lives. To address this issue, researchers recommend that organizations provide clear and accessible information about their data collection and processing practices, as well as the algorithms they use. Additionally, organizations should engage in open dialogue with stakeholders, including individuals, community groups, and policymakers, to foster trust and accountability.

5. Conclusion

The increasing use of data and science engineering in various domains of society has brought with it a range of ethical challenges that need to be addressed. Through a comprehensive literature review of 18 studies, this research identified and analyzed the key ethical issues in data and science engineering, the impact of these issues on society, and proposed solutions for addressing them. The findings of this study revealed that the most prominent ethical issues in data and science engineering are privacy breaches, algorithmic biases, and lack of transparency. These issues have a significant impact on society, particularly on vulnerable groups such as minorities and marginalized communities. The study also identified the need for greater regulation and accountability in the use of data and science engineering.

The proposed solutions for addressing these ethical issues include the adoption of ethical guidelines and standards, increased transparency and accountability, the development of fair and unbiased algorithms, and the empowerment of individuals to control their own data. These solutions require collaboration between policymakers, industry professionals, and society at large. The feasibility of implementing these solutions depends on several factors, including political will, technological advancements, and the willingness of stakeholders to take responsibility for their actions. Despite these challenges, the potential benefits of addressing these ethical issues are significant, including increased trust and credibility in data and science engineering, improved outcomes for individuals and society, and the avoidance of harmful impacts.

In conclusion, this research highlights the urgent need for greater ethical considerations in data and science engineering. The identified ethical issues and proposed solutions can serve as a guide for policymakers, industry professionals, and society at large to ensure that the use of data and science engineering is responsible, ethical, and beneficial for all. It is essential to balance the potential benefits of these technologies with the need to ensure that they are used ethically and responsibly.

Compliance with ethical standards

Acknowledgments

We would like to express our deepest gratitude to all those who have contributed to this research paper. Firstly, we would like to thank all the authors whose studies were reviewed in this paper. Their research has provided valuable insights into the ethical challenges that arise from the use of data and science engineering, and their contributions are greatly appreciated. We extend our gratitude to our academic advisors and mentors, who provided us with valuable guidance and feedback throughout the research process. Their expertise and support were instrumental in ensuring the quality of this paper.

Disclosure of conflict of interest




There is no conflict of interest in this study.

References

- [1] Floridi, Luciano & Taddeo, Mariarosaria. (2016). What is data ethics?. *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, 374. 20160360. 10.1098/rsta.2016.0360.
- [2] Star, S. L., & Ruhleder, K. (1994, October). Steps towards an ecology of infrastructure: complex problems in design and access for large-scale collaborative systems. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp. 253-264).
- [3] Floridi, L., & Illari, P. (Eds.). (2014). *The philosophy of information quality* (Vol. 358). Cham: Springer.
- [4] Levin, N., Leonelli, S., Weckowska, D., Castle, D., & Dupré, J. (2016). How do scientists define openness? Exploring the relationship between open science policies and research practice. *Bulletin of science, technology & society*, 36(2), 128-141.
- [5] Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. *Social studies of science*, 41(5), 667-690
- [6] Brunton, S. L., & Kutz, J. N. (2022). *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press
- [7] Egger, R., Neuburger, L., & Mattuzzi, M. (2022). Data Science and Ethical Issues: Between Knowledge Gain and Ethical Responsibility. In *Applied Data Science in Tourism: Interdisciplinary Approaches, Methodologies, and Applications* (pp. 51-66). Cham: Springer International Publishing
- [8] Hammouchi, H., Cherqi, O., Mezzour, G., Ghogho, M., & El Koutbi, M. (2019). Digging deeper into data breaches: An exploratory data analysis of hacking breaches over time. *Procedia Computer Science*, 151, 1004-1009.
- [9] Ayereby, M. P. M. (2018). *Overcoming data breaches and human factors in minimizing threats to cyber-security ecosystems* (Doctoral dissertation, Walden University).
- [10] Lee, N. T., Resnick, P., & Barton, G. (2019). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. Brookings Institute: Washington, DC, USA, 2.

- [11] Hooker, S. (2021). Moving beyond “algorithmic bias is a data problem”. *Patterns*, 2(4), 100241.
- [12] Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D’Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, 60, 102387. <https://doi.org/10.1016/j.ijinfomgt.2021.10238>
- [13] Kocak, B., Yardimci, A. H., Yuzkan, S., Keles, A., Altun, O., Bulut, E., ... & Okumus, A. A. (2022). Transparency in Artificial Intelligence Research: a Systematic Review of Availability Items Related to Open Science in Radiology and Nuclear Medicine. *Academic Radiology*
- [14] Callaghan, S. (2020). On the importance of data transparency. *Patterns*, 1(4).
- [15] Hosseini, M., Wiczorek, M., & Gordijn, B. (2022). Ethical issues in social science research employing big data. *Science and Engineering Ethics*, 28(3), 29.
- [16] Da Bormida, M. (2021). The Big Data World: Benefits, Threats and Ethical Challenges. In *Ethical Issues in Covert, Security and Surveillance Research* (Vol. 8, pp. 71-91). Emerald Publishing Limited.
- [17] Amon, J. J., Baral, S. D., Beyrer, C., & Kass, N. (2012). Human rights research and ethics review: protecting individuals or protecting the state?.
- [18] Da Bormida, M. (2021). The Big Data World: Benefits, Threats and Ethical Challenges. In *Ethical Issues in Covert, Security and Surveillance Research* (Vol. 8, pp. 71-91). Emerald Publishing Limited.

Author’s short Biography

	<p>Saloni Saloni is a chemical engineer with a bachelor's degree from KU Leuven in Belgium. Her research interests include chemical engineering, civil engineering, computing, AI, data engineering, data modelling, and machine learning algorithms. With her background in chemical engineering, Saloni has a solid foundation in scientific and technical knowledge, which has allowed her to apply her expertise to various projects. Additionally, Saloni is passionate about exploring new fields of research and has pursued projects in areas that intrigue her, such as data engineering and machine learning algorithms.</p>
	<p>Shubham Shubham is a software engineer and independent researcher with a bachelor's degree in civil engineering from Technion - Israel Institute of Technology and a master's degree in building construction from Georgia Institute of Technology. Their research interests include data engineering, data modelling, machine learning, AI, and computing in construction and infrastructure. As a software engineer, Shubham has worked on numerous projects, developing their technical skills and gaining practical experience in the field. In addition, Shubham has pursued independent research projects related to their interests, contributing to the advancement of knowledge in these areas.</p>
	<p>Sidra-Tul-Muntaha is a highly accomplished professional researcher with a Bachelor's degree in English from Fatima Jinnah Women's University Rawalpindi and a degree in Health Sciences from the University of People America. With almost six years of research experience, she has worked on numerous research papers and articles in reputed journals and has presented her work at national conferences.</p>