

Big Data: Ethics and Governance in Business Analytics

Dr Triveni P.

Professor-Finance
triveni@msrim.org

Sanjana R

Student
sanjanar241136@msrim.org

Yoganandachari D V

Student
yoganandacharidv241150@msrim.org

Submitted to 10th Global Business Environment (GBE) International Conference
“International Business Diplomacy, Global Trade, and Economic Development in
Developing Countries”

Organised by IMI Bhubaneswar and EGADE Business School Mexico City, Mexico.

Article Type: Research Article

Article Citation: Dr Triveni P., Sanjana R, Yoganandachari D V, Big Data: Ethics and Governance in Business Analytics. M.S. Ramaiah Management Review. 2025; 16(02), 114-123. DOI: 10.52184/msmr.v16i02.113

Received date: October 15, 2024

Accepted date: December 15, 2024

***Author for correspondence:** Dr Triveni P.  triveni@msrim.org

 Professor-Finance

Abstract

With the ever-increasing focus on artificial intelligence, there is an exponential growth of data generation from government, private and public. There is increase in the number of data generating devices as well, such as smart devices, wearable devices, and the everyday used and seen sensor devices around us. It is essential to understand the governance related to the data and its privacy along with the ethical issues to surrounding the usage of data. Data breach amongst the individuals and organizations is a major challenge in today's business world. The data which is a powerful resource for decision making and to scale the businesses is to be validated amongst the challenges of collecting storage, usage and the rights associated. The field of data ethics explores these questions and offers guiding principles for business professionals who handle data, there exists a significant risk of data exploitation for both individuals and companies. Business intelligence, big data and business analytics is the buzz in this new era of business data. Big data analytics and business analytics carry a lot of similarities: they both bite off some data, chew it up and spit it out as some new form of cohesive, useful information supporting the decision making of an organization. In the same vein, business analytics is very human-focused, while big data analytics requires too much processing and focused approach to be adopted to analyse the data.

This study revolves around the 5 principles of data ethics provided by Harvard business school(online) and the governance of big data by Kate Prochorik , data intelligence researcher. The five Harvard data ethics are: Ownership, transparency, privacy, Intention and outcomes.⁹ Even when intentions are good, the outcome of data analysis can cause inadvertent harm to individuals or groups of people. Unfortunately, instead of reaping instant benefits, business users have realized that in spite of data governance frameworks which exists they have to execute an extended set of data analytics to solve challenges connected with multiple data formats and security. There are 4 big data governance components: people, polices procedures standards, technology and infrastructure and support and optimization. According to Kate there are three data governance levels: strictly governed, loosely governed and nongoverned data which we aim to delve deep into in this research. The study aims to provide a grasp of the state of big data ethics and its governance in large organizations and corporations. This study aims to uncover the loop holes in the existing frame works in data privacy, security and transparency. This study will give an overview of the co-operation of different functions of an organization, the current systems in place for ethics violation and steps involved in whole data lifecycle used for data analytics. In the end, the study offers the state of ethics and governance of big data in business analytics and the policies in place and provide suggestions if any.

Keywords: Artificial Intelligence (AI), Big data, Ethics, Privacy, Breach, Governance.

Introduction:

Greater is the volume of data through rapid technological advancement. Therefore, there comes the unprecedented amount of knowledge to be available for any business enterprise. Thus, in big data analytics, ability to process huge volumes of structured as well as unstructured data has enabled it as a powerful tool for innovations and decision-making. As the size, speed, and complexity of the data continue to expand, the need for the ethical and governance issues on the process of collecting, storing, analyzing, and utilizing such data is also felt. Serious ethical dilemmas exist within big data analytics

with regards to privacy issues: Big data analytics involves collection and processing of vast amounts of personal information, giving rise to a lot of critical concerns about privacy. Clear consent must be obtained, and business should ensure tight security control and adherence to data protection law as required by GDPR and CCPA. The impact of financial loss, damage to reputation, and legal liabilities can be pretty harsh on business and individuals if the breach occurs. Major ethical issues are algorithm bias. Algorithms tend to embed biases learned at training and may thus provide outcome discrimination. Bias mitigation at data and algorithmic levels, therefore,

involves careful data curation, insightful algorithm design, as well as over-time monitoring for fairness and equity. They are the very heart of ethics of big data: transparency and accountability. In that sense, organizations ought to be transparent about practices related to their data collections, storage, and use practices. Such organizations should then be answerable for adverse ramifications of such choices based on the data. This will require openness over capabilities and limitations of models worked upon, uncertainties in statements of certain facts and not only tolerating but welcoming proper critiques. Organizations would thus require an end-to-end data governance framework for dealing with all such ethics-related issues. It will embrace data quality, security, privacy, and ethics policies and procedures that are in place; second, it will identify distinct roles and responsibilities associated with managing data—that is, who is going to do what. Thus, data will be dealt with very high responsibility and ethics. Data governance is the base to dealing with data in every phase of its life cycle efficiently and effectively. Key issues under the governance category are: data quality, data security, data privacy, and access to data. Data quality will ensure proper and accurate analysis. Organizations must carry out data quality checks to ensure accuracy, completeness, consistency, and relevance of data. Data security deals with matters related to unauthorized access, use, disclosure, disruption, modification, or destruction of sensitive data. Data privacy safeguards personal information and observes corresponding laws and regulations. Access controls relating to data must be established to restrict its access only to authorized personnel who would otherwise misuse it. Organizational leadership and, more importantly, executive support would

also make a big difference in data-driven initiatives. Organizations should adopt a culture that focuses on the quality, security, and privacy of data. The investments could be in tools like data management training for employees and proper setting of guidelines for data governance.

Review of Literature:

Big Data technologies have rapidly expanded to a sizeable level, bringing critical concerns with them into ethical and governance across all disciplines that fall into the folds, starting from health and public health all through urban studies up to human resource management. Organizations now take crucial decisions, given reliance on data. Consequently, ethics must have their supportive strong governance structure with protective safeguard on human rights. The question, however, is not unique in health only but extends into questions of ownership and thus use of personal data, moving on to governance queries in such urban studies (Floridi & Taddeo, 2021). Shamim et al. (2020) explain how data literacy would address the problem of miscommunication and help with proper, ethical data usage. As argued, only through developing skills that can adequately surpass Big Data complexities can one achieve governance that is ethically managed. This view is within line in the broader Big Data discussion on ethics, which makes informed citizenry actionable to better protect rights as well as responsibly interact in this practice of data. Transparency and accountability are the fundamental components of ethical governance frameworks. According to Mikalef et al. (2020), information governance plays a significant role in Big Data analytics-driven

innovation. Proper ethical rules should govern data practices. The same concept is expressed by Stahl and Wright (2018), who support the responsible research and innovation principles used in AI and Big Data technology development. Ensuring transparency in the practice of data collection, storage, and sharing is essential to build up trust among stakeholders. According to Kumar et al. (2022), understanding the interplay of these technologies with ethical considerations would be critical for developing comprehensive governance mechanisms. The interaction of Big Data ethics with global issues, such as inequality and access to data, is an area that deserves further exploration. Critical questions on the political economy of Big Data in development contexts have been posed by Mann (2017), who argued that future research should look into how ethical governance can deal with disparities in data access and usage.

Research Gap:

Much work has been carried out so far on big data ethics and governance matters in business analytics, but yet there is an unknown gap in how emerging economy organizations face their respective issues.

Statement of Problem:

Rapid big data development has brought a package of opportunities and challenges. It presents some useful insights about the system while raising pertinent issues concerning privacy, security, and ethics. All these have provided a complex arena of regulation over data, an ethical context, and related technological support. For

organizational realization, data governance policies must adopt the framework under which data quality, security, privacy, and transparency are pursued. By addressing these challenges, organizations may harness the power of big data while minimizing risks, ensuring ethical and responsible use of data.

Objectives:

1. To understand Ethical and governance challenges of big data in Business Analytics.
2. To analyse the existing loop holes and attempt to provide solution to them.
3. To study the regulatory framework of countries like USA, UK, China and India.

Research methodology:

Research Methodology: Descriptive Research

Source of Data: Secondary data.

Analysis:

Objective 1: To understand Ethical and governance challenges of big data in Business Analytics.

The rapid advance of technology has drastically increased the production of data, therefore making available opportunities for insights by enterprises. However, this data-centric transformation also poses steep ethical and governance dilemmas. As institutions continually rely on big data analytics for imperative decision-making activities, they need to address the

challenges of privacy, security, bias, and transparency.

Ethical challenges:

- **Data breaches:** The dependence on technology has led technology firms to gather enormous volumes of private data. The information is often used by the parties to tailor promotional messages and direct user behaviour, increasing serious issues about privacy. In addition, there is a rising possibility of suffering from data breaches due to more sophisticated cyberattacks being developed. A data breach may compromise sensitive personal data, and identity theft, financial loss, and reputational damage could occur when it happens. One of the example for this is the Cambridge Analytica scandal which used data from Facebook to predict what political party was favoured in Britain.
- **Algorithmic bias:** The increasing use of AI has triggered more concern over bias and fairness. For example, facial recognition is primarily applied to watching and securing the people, and this has been regarded as biased towards other races and genders. In some cases, these systems misidentify some racial or ethnic groups of people, leading to unfair results. AI hiring machines have proven to discriminate against certain groups of people based on the likes of gender or race, which extends the present inequalities.

One of the main examples is when Amazon implemented AI software for hiring in the year 2014 and it had shown discrimination towards hiring of females.

- **Transparency and accountability:** Being transparent will help build trust from

customers and others working with the organization. Additionally, organizations should be accountable for all negative impacts of their data-informed decisions. For example, Google has been criticized about its data collection, most especially about how users are being tracked and the ads that are targeted to users. Although it is true that these methods enhance user experience, they have also made people worried about privacy and the possibility of manipulation.

Governance challenges:

- **Data quality:** Data quality is a critical aspect of analysis and decision-making. Issues in data quality, such as missing values, inconsistencies, and errors, may give rise to incorrect insights and poor results. Organizations need to implement robust data quality management practices, which include data cleaning, validation, and standardization. Identifying errors, filling up missing values, and ensuring data consistency across different sources fall under this practice.
- **Data security:** Protecting sensitive information from cyber attacks is of paramount importance for any enterprise. It has to assure the building of an integrated security framework. It includes encryption for protection against unauthorized access, strict access control which denies access to only authorized individuals, regular security auditing to identify vulnerabilities and fix them, and strong incident response strategies that provide robust damage management in case a breach happens.
- **Data privacy:** Properly safeguard information related to an individual and also data privacy law is imperative. Organizations must receive explicit permission from a user before collecting any of

their personal data for further processing purposes. The organization must also ensure appropriate measures to safeguard the data from unauthorized access and misuse. Moreover, organizations must meet the regulation of the related data protection laws such as GDPR, CCPA, etc., which call very stringent regulations on the collection, storage, and other processes related to personal data.

- **Data access and sharing:** Access rules and procedures should therefore be defined in order to prevent unauthorized access and its misuse. Organizations must, therefore use role-based access controls where the right people have access to sensitive data. Data sharing agreements should also be made on managing how data is shared outside of the organization.
- **Data retention and disposal:** Organizations must have clearly defined policies on how to deal with the retention and disposal of data in an effort to minimize breaches of data and infractions. Data retention policies should specify for how long data should be retained and under what circumstances. Data disposal policies should detail the processes for securely deleting or destroying data if determined to be no longer useful.

Objective 2: To analyse the existing loop holes and attempt to provide solution to them.

Although considerable strides have been made toward the mitigation of ethical and governance issues arising from big data, some loopholes persist. They include the lack of transparency in the collection and use of data, algorithmic bias, and difficulties

in implementing data privacy regulations. It is in such areas that organizations need to pay greater attention to data ethics, create strong governance frameworks, and stay informed about emerging technologies and regulations to mitigate the risk factors involved.

- **Lack of transparency and accountability:** Many organizations do not inform people about their data operations. This makes it hard to rationalize usage of data. The lack of clarity may eventually lead to misuse of data and loss of trust. To address this issue, organizations are told to implement comprehensive data governance frameworks that provide maximized transparency and accountability. This would include coming up with clear data privacy policies, the necessity for regular audit processes, as well as the provision of user self-regulation over private data.
- **Insufficient data security measures:** Poor security policies can compromise the organization to several cyberattacks and data theft affecting the confidential information used in organizations. Organizations are therefore advised to adopt systematic security approaches like encryption, mechanisms of access control, security audits among other techniques protect data. Incident response plans development can also make the consequences of security breaches minimal. Moreover, regular employee training about best practice data security can increase general security awareness.
- **Rapid advancements in technology:** Technology innovation is developing at very rapid rates; these innovations in the technique of data governance outstrip the establishment of efficient

data frameworks, which creates new challenges and risks. For that reason, organizations must remain informed regarding current trends and their possible implications for data governance. Applying new technologies, such as artificial intelligence and machine learning, can improve data governance methods. Furthermore, collaboration with industry experts and participating in discussions related to the subject matter will help organizations stay updated.

- **Inconsistent Data Quality Standards:** Different departments or teams of an organization can have different data quality standards that lead to

inconsistencies and errors of analysis. Organizations should set the standards of high-quality data and enforce it. Data cleansing and validation involves error identification, which is then corrected. Techniques of data profiling can be described in describing characteristics of data and noting probable issues. In addition, data governance teams can be formed within organizations to monitor data quality and ensure that the standards are maintained. This means that organizations can ensure accuracy and reliability in their analytics and decision-making processes by prioritizing data quality.

Objective 3: To study the regulatory framework of countries like USA, UK, China and India.

USA Regulatory frame work

HIPAA: Health Insurance Portability and Accountability Act	Can result in fines of up to \$250,000 and imprisonment for up to 10 years in cases of intentional misuse.
Gramm-Leach-Bliley Act (GLBA): (12 November 1999)	Civil penalties of up to \$100,000 per violation for financial institutions. Personal liability for executives, with fines up to \$10,000. Criminal charges for severe violations, including imprisonment for up to 5 years.
Children’s Online Privacy Protection Act (COPPA): 21 April 2000	Civil penalties of up to \$50,120 per violation as of the latest FTC adjustments for inflation. Ongoing violations can result in significant cumulative fines and mandatory compliance orders.
California Consumer Privacy Act (CCPA) 1 January 2020	Civil penalties of up to \$7,500 per intentional violation and \$2,500 for unintentional violations. Statutory damages of \$100 to \$750 per consumer, per incident, in cases of data breaches.

China Regulatory frame work

PIPL (Personal Information Protection Law, 20 August 2021)	50 million RMB or 5% of annual revenue
DSL (Data Security Law, 1 Sep 2021)	10 million RMB

Cybersecurity Law, (7 Nov 2016)	1 million RMB (general); 5% of revenue Suspension of operations, blacklisting for critical violations.
------------------------------------	--

UK Regulatory frame work

	Aligns with the UK GDPR fines (as listed above). Additional penalties for mishandling sensitive data related to criminal investigations or national security.
Data Protection Act 2018	
Privacy and Electronic Communications Regulations (PECR)	Fines up to £500,000 for breaches, issued by the Information Commissioner's Office (ICO). Recent examples include penalties for unsolicited marketing calls, failure to obtain cookie consent, or spam emails
Investigatory Powers Act 2016 (IPA)	Criminal Charges: Unauthorized interception or misuse of surveillance powers can lead to fines and imprisonment. Potential disqualification of organizations from government contracts or operations if they fail to comply with lawful requirements. Specific penalties are case-dependent, with courts issuing judgments for violations

India Regulatory frame work Regulatory frame work

	Section 43A: Compensation of up to ₹5 crores for negligence in data protection resulting in harm.
Information Technology Act, 2000	Section 66: Imprisonment of up to 3 years or fines up to ₹5 lakh for hacking and unauthorized access.
The Aadhaar (Targeted Delivery of Financial and Other Subsidies) Act, 2016	Unauthorized Sharing or Misuse of Aadhaar Data: Imprisonment up to 3 years and fines up to ₹10,000 (or ₹1 lakh for companies). Accessing Aadhaar Data Without Authorization: Imprisonment of up to 10 years for severe breaches of confidentiality. Civil Liability: Damages may be imposed on organizations responsible for breaches
National Data Governance Framework Policy (NDGFP), 2023	NDGFP does not impose specific penalties but aligns with the principles of related laws (like the DPD Act and IT Act). Penalties are likely to be enforced through linked acts like the IT Act for misuse of non-personal data.

Conclusion:

Rapid technological progress has led to exponential growth in the generation and consumption of data. This, in turn, has presented unparalleled opportunities for businesses to make critical decisions with greater insights. Although big data is an immense opportunity for innovation and societal development, it also throws

enormous ethical and governance challenges. Hence, with organizations heavily dependent on big data analytics for critical decision-making, the time has come to address issues such as privacy, security, bias, and transparency. As a result, there should be a comprehensive approach to the data governance of organizations who embrace big data with minimal risk. This means clarity into data governance policies

and all security measures with an essence for data privacy. More essentially, there should be the ethos in data ethics where employees find out about the ethical expectations in data practice. The policymakers and regulators will define the future of big data. It can be assured for the government that organizations will have been responsible and ethical by their use of data based on the laws to make effective data protection. Flows are international, and the requirement for standard harmonization regarding data privacy and security shouldered international cooperation. Such knowledge by organizations about ethical and governance issues would mean an enlightened decision through use of data towards betterment. Technologies are fast moving, other than some associated challenges and opportunities created with big data.

Recommendations:

In dealing with the issues related to big data, there should be data privacy and security as a priority, have data governance frameworks in place and establish a culture of data ethics. Organizational-level data privacy can be addressed by explicit consent, not collecting much data, and ensuring robust security. Data quality standards should be set up in an organization; periodical audits should be performed and stakeholder engagement in discussions over matters that are clear, accountable, and transparent should be encouraged.

References

1. Mittelstadt, B.D., Floridi, L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng*

2. Shamim, Saqib., Zeng, J., Khan, Zaheer., & Zia, Najam Ul. (2020). Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms. *Technological Forecasting and Social Change* . <http://doi.org/10.1016/J.TECHFORE.2020.120315>
3. Mikalef, Patrick., Boura, Maria., Lekakos, George., & Krogstie, J. (2020). The role of information governance in big data analytics driven innovation. *Inf. Manag.* , 57 , 103361 . <http://doi.org/10.1016/j.im.2020.103361>
4. B. C. Stahl and D. Wright, "Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation," in *IEEE Security & Privacy*, vol. 16, no. 3, pp. 26-33, May/June 2018, doi: 10.1109/MSP.2018.2701164.
5. Kumar, S., Sharma, D., Rao, S. *et al.* Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. *Ann Oper Res* (2022). <https://doi.org/10.1007/s10479-021-04410-8>
6. Mann, Laura. (2017). Left to Other Peoples' Devices? A Political Economy Perspective on the Big Data Revolution in Development. *Development and Change* , 49 , 3-36 . <http://doi.org/10.1111/DECH.12347>
7. https://link.springer.com/chapter/10.1007/978-3-031-51063-2_5
8. <https://www.ibe.org.uk/resource/business-ethics-and-big-data.html>
9. <https://online.hbs.edu/blog/post/data-ethics>
10. <https://www.itransition.com/blog/big-data-governance#:~:text=A%20big%20data%20governance%20>

- strategy%20requil. r e s % 2 0 a % 2 0 professional,data%20scientists%2C%20 data%20owners%2C%20data%20 stewards%2C%20or%20others
12. <https://www.selecthub.com/big-data-analytics/big-data-business-analytics/>
 13. Mittelstadt, B.D., Floridi, L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng Ethics* **22**, 303–341 (2016). <https://doi.org/10.1007/s11948-015-9652-2>
 14. Shamim, Saqib., Zeng, J., Khan, Zaheer., & Zia, Najam Ul. (2020). Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms. *Technological Forecasting and Social Change* . <http://doi.org/10.1016/J.TECHFORE.2020.120315>
 15. Mikalef, Patrick., Boura, Maria., Lekakos, George., & Krogstie, J.. (2020). The role of information governance in big data analytics driven innovation. *Inf. Manag.*, *57* , 103361 . <http://doi.org/10.1016/j.im.2020.103361>
 16. B. C. Stahl and D. Wright, “Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation,” in *IEEE Security & Privacy*, vol. 16, no. 3, pp. 26-33, May/June 2018, doi: 10.1109/MSP.2018.2701164.
 17. Kumar, S., Sharma, D., Rao, S. *et al.* Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. *Ann Oper Res* (2022). <https://doi.org/10.1007/s10479-021-04410-8>
 18. Mann, Laura. (2017). Left to Other Peoples’ Devices? A Political Economy Perspective on the Big Data Revolution in Development. *Development and Change*, *49* , 3-36 . <http://doi.org/10.1111/DECH.12347>